OM of: US-10-031-904-8 to: N_Geneseq_032802:* out_format : pfs

Date: Oct 9, 2002 6:42 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10031904/runat_09102002_084257_23681/app_query.fasta_1.232
-DB=N_Geneseq_032802 -QFMT=fastap -SUFFIX=p2n.rng -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10031904_@CGN1_1_481 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-10-031-904-8
Query length: 174
Database: N_Geneseq_032802:*
Database sequences: 1736436
Database length: 858457221
Search time (sec): 175.470000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	! Documentation ..			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAF58602	+	946.00	1870.05	8.2e-96	6				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:AAS64474	+	650.50	1253.09	1.9e-61	7				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:AAN91477	+	649.50	1252.37	2.1e-61	69				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:AAQ11642	+	649.50	1252.37	2.1e-61	69				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:AAZ38150	+	649.50	1252.37	2.1e-61	69				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAI58380	+	649.50	1251.82	2.2e-61	7				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:ABA09026	+	636.50	1226.31	5.9e-60	7				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAI60166	+	636.50	1226.31	5.9e-60	7				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:AAV53262	+	635.00	1250.09	2.8e-61	59				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:AAV53269	+	635.00	1250.09	2.8e-61	59				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:AAV53270	+	635.00	1250.09	2.8e-61	59				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:AAV53271	+	635.00	1250.09	2.8e-61	59				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:AAV53272	+	635.00	1250.09	2.8e-61	59				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:AAV53273	+	635.00	1250.09	2.8e-61	59				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:AAQ41867	+	634.50	1222.44	9.6e-60	69				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:AAS64473	+	513.00	989.83	8.7e-47	2				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:AAS64289	+	487.00	937.94	6.8e-44	2				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:AAS64290	+	458.00	867.39	5.8e-40	9				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:AAQ11643	+	442.00	840.99	1.7e-38	54				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:AAZ38151	+	442.00	840.99	1.7e-38	54				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:ABA01504	+	308.00	592.31	1.2e-24	9				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:AAT17595	+	289.00	552.59	2.0e-22	11				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:AAT17597	+	289.00	552.59	2.0e-22	11				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:AAQ14919	+	289.00	551.59	2.2e-22	12				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:AAT90306	+	289.00	549.74	2.8e-22	14				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:AAQ99106	+	289.00	549.36	3.0e-22	15				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:AAT03339	+	289.00	549.36	3.0e-22	15				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:AAQ10864	+	289.00	549.24	3.0e-22	15				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:AAQ14915	+	289.00	548.48	3.3e-22	16				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAD20355	+	289.00	547.31	3.9e-22	1				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAD08170	+	289.00	547.31	3.9e-22	1				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:AAT46065	+	289.00	547.14	4.0e-22	18				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:AAQ14916	+	289.00	546.51	4.3e-22	19				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAH34933	+	289.00	541.04	8.7e-22	3				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAF18270	+	289.00	541.02	8.7e-22	33				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:AAT17598	+	287.00	548.60	3.3e-22	11				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:AAT17596	+	287.00	548.60	3.3e-22	11				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:AAV08935	+	272.50	519.55	1.4e-20	11				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:AAX87915	+	272.50	516.69	2.0e-20	14				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:AAN70047	+	272.50	513.45	3.0e-20	20				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:AAX87914	+	272.50	513.00	3.2e-20					
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:AAN91043	+	272.50	512.93	3.2e-20					
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:AAQ79863	+	272.50	512.93	3.2e-20					
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:AAN70048	+	272.50	512.34	3.4e-20					
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:AAN91044	+	272.50	512.34	3.4e-20					

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAF58602

seq_documentation_block:
ID AAF58602 standard; cDNA; 627 BP.
XX
AC AAF58602;
XX
DT 24-APR-2001 (first entry)
XX
DE Human RECAP polynucleotide, SEQ ID NO: 30.
XX
KW Human; RECAP; receptors and associated proteins; cerebroprotective;
KW nootropic; neuroprotective; anticonvulsant; antiparkinsonian; anti-HIV;
KW antidiabetic; immunostimulant; immunomodulator; antiinflammatory;
KW antithyroid; immunosuppressive; nephrotropic; antigout; thyromimetic;
KW cytostatic; antibacterial; virucide; fungicide; protozoacide;
KW antiarteriosclerotic; hepatotropic; gene therapy; infection; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200107612-A2.
XX
PD 01-FEB-2001.
XX
PF 21-JUL-2000; 2000WO-US20035.
XX
PR 21-JUL-1999; 99US-0145232.
PR 07-OCT-1999; 99US-0158578.
PR 12-NOV-1999; 99US-0165192.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Au-Young J, Bandman O, Tang YT, Yue H, Azimzai Y, Burford N;
PI Baughn MR, Lu DAM, Hillman JL, Patterson C, Lal P;
XX
DR WPI; 2001-168554/17.
DR P-PSDB; AAB68878.
XX
PT Novel receptors and associated proteins for diagnosis and treatment of
PT neurological disorders, immunological disorders including autoimmune/
PT inflammatory disorders and cell proliferative disorders such as cancer
PT -
XX
PS Claim 5; Page 120; 128pp; English.
XX
CC The present sequence encodes a human RECAP (receptors and associated
CC proteins) polypeptide. RECAP polynucleotides and polypeptides are useful
CC in the diagnosis, treatment and prevention of neurological disorders
CC such as stroke, Alzheimer's disease, Pick's disease, Huntington's
CC disease, dementia, Parkinson's disease, Downs's syndrome, amyotrophic
CC lateral sclerosis, multiple sclerosis, bacterial and viral meningitis,
CC CJD (Creutzfeldt-Jakob disease), GSS (Gerstmann -Straussler-Scheinker
CC syndrome); immunological disorders, including autoimmune/inflammatory
CC disorders such as AIDS, DiGeorge's syndrome, severe combined
CC immunodeficiency disease (SCID), Chediak-Higashi syndrome, Cushing's
CC disease, Addison's disease, autoimmune thyroiditis, Crohn's disease,
CC diabetes mellitus, Good pasture's syndrome, gout, Grave's diseases,
CC Hashimoto's thyroiditis, Sjogren's syndrome, Werner's syndrome, viral,
CC bacterial, fungal, parasitic, protozoal, and helminthic infections; and
CC cell proliferation disorders such as arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis and cancer.
XX
SQ Sequence 627 BP; 142 A; 173 C; 128 G; 184 T; 0 other;

alignment_scores:
Quality: 946.00 Length: 174

```
            Ratio:  5.437              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-10-031-904-8 x AAF58602   ..

 Align seg 1/1  to: AAF58602  from: 1  to: 627

       1 MetAlaProProValArgLeuGluArgProPheProSerArgArgPhePr 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      87 ATGGCGCCTCCCGTCCGTCTCGAGCGTCCCTTTCCTTCCCGGCGCTTTCC 136

      17 oGlyLeuLeuLeuAlaAlaLeuValLeuLeuLeuSerSerPheSerAspG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     137 TGGGTTGCTTCTGGCGGCCCTGGTGTTGCTGCTGTCCTCCTTCTCCGATC 186

      34 lnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     187 AATGCAATGTCCCGGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACT 236

      51 AspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgPr 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     237 GATGACTTTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCC 286

      67 oGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerValT 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     287 TGGTTATTCCGGAAGACCGTTTTCTATCATCTGCCTAAAAAACTCAGTCT 336

      84 rpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProPro 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     337 GGACAAGTGCTAAGGACAAGTGCAAACGTAAATCATGTCGTAATCCTCCA 386

     101 AspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlySe 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     387 GATCCTGTGAATGGCATGGCACATGTGATCAAAGACATCCAGTTCGGATC 436

     117 rGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSerS 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     437 CCAAATTAAATATTCTTGTCCTAAAGGATACCGACTCATTGGTTCCTCGT 486

     134 erAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysThr 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     487 CTGCCACATGCATCATCTCAGGCAACACTGTCATTTGGGATAATAAAACA 536

     151 ProValCysAspSerGluLeuLysTyrAlaPheLeuPheLeuLeuProIl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     537 CCTGTTTGTGACAGTGAGTTGAAATATGCATTCCTATTTCTTTTACCGAT 586

     167 eHisSerAsnPheSerLeuGlu 174
         ||||||||||||||||||||||
     587 ACATTCTAATTTTTCTCTGGAA 608

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:AAS64474

seq_documentation_block:
ID   AAS64474 standard; cDNA; 7821 BP.
XX
AC   AAS64474;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #278.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG00287.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 278; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 7821 BP; 1986 A; 1904 C; 1883 G; 2046 T; 2 other;


alignment_scores:
            Quality:  650.50              Length:    149
              Ratio:  4.854                 Gaps:      2
 Percent Similarity:  89.933   Percent Identity: 81.208

alignment_block:
 US-10-031-904-8 x AAS64474   ..

 Align seg 1/1  to: AAS64474  from: 1  to: 7821

       9 ArgProPheProSerArgArgPhe......ProGlyLeuLeuLeuAlaAl 23
         :::|||   |||:::      |||         |||   |||||||||::
     497 CAGCCGGCGCCCGGTCTCCCCTTCTGCTGCGGAGGATCCCTGCTGGCGGT 546

      23 aLeuValLeuLeu...LeuSerSerPheSerAspGlnCysAsnValProG 39
         ::::|||||||||   |||               ||||||||::::||||
     547 CGTGGTGCTGCTCGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAG 596

      39 luTrpLeuProPheAlaArgProThrAsnLeuThrAspAspPheGluPhe 55
         ||||||||||||||||||||||||||||||||||:::|||||||||
     597 AATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTT 646

      56 ProIleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyAr 72
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     647 CCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAG 696

      72 gProPheSerIleIleCysLeuLysAsnSerValTrpThrSerAlaLysA 89
         ||||||||||||||||||||||||||||||||||:::|||||||
     697 ACCGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACTGGTGCTAAGG 746
```

```
 89 spLysCysLysArgLysSerCysArgAsnProProAspProValAsnGly 105
    ||:::|||:::||||||||||||||||||||||||||||||||||||||||
747 ACAGGTGCAGACGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGC 796

106 MetAlaHisValIleLysAspIleGlnPheGlySerGlnIleLysTyrSe 122
    |||:::||||||||||||   ||||||||||||||||||||||||||||
797 ATGGTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAAATTAAATATTC 846

122 rCysProLysGlyTyrArgLeuIleGlySerSerSerAlaThrCysIleI 139
    ||||   ||||||||||||||||||||||||||||||||||||||||||
847 TTGTACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCA 896

139 leSerGlyAsnThrValIleTrpAspAsnLysThrProValCysAsp 154
    ||||||||:::|||||||||||||||||:::|||||||:::|||||||
897 TCTCAGGTGATACTGTCATTTGGGATAATGAAACACCTATTTGTGAC 943
```

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:AAN91477

seq_documentation_block:
ID AAN91477 standard; DNA; 6951 BP.
XX
AC AAN91477;
XX
DT 22-FEB-1990 (first entry)
XX
DE CR1 protein DNA.
XX
KW Complement; cofactor.
XX
OS Homo sapiens (human).
XX

FH	Key	Location/Qualifiers
FT	sig_peptide	28..147
FT		/*tag= a
FT	CDS	1..6951
FT		/*tag= b
FT	CDS	28..6147
FT		/*tag= c
FT		/note="CR1 protein"
FT	CDS	28..1533
FT		/*tag= d
FT		/note="CR1 protein"

PN WO8909220-A.
XX
PD 05-OCT-1989.
XX
PF 31-MAR-1989; 89WO-US01358.
XX
PR 01-APR-1988; 88US-0176532.
XX
PA (TCEL-) T CELL SCIENCES INC.
PA (UYJO) THE JOHNS HOPKINS UNIVERSITY.
PA (BRIG) THE BRIGHAM AND WOMENS HOSPITAL.
XX
PI Fearon DT, Klickstein LB, Wong W, Carson G, Concino MF, Makrides SC;
XX
DR WPI; 1989-309498/42.
DR P-PSDB; AAP92219.
XX
PT New nucleic acid sequences encoding new CR1 protein - and its fragment,
PT for diagnosis and control of complement-related immune defects,
PT inflammation, myocardial infarct, etc
XX
PS Claim 1; fig. 1; 191pp; English.
XX
CC This is full-length CR1 protein, and shortened forms are new, lacking
CC the transmembrane region. The proteins and fragments bind C3b and/or
CC C4b, have cofactor activity and inhibit C3 and C5 convertase activity.
CC They are useful in diagnosing and treating immune disorders, and prevent
CC perfusion injury.
XX
SQ Sequence 6951 BP; 1802 A; 1680 C; 1660 G; 1809 T; 0 other;

alignment_scores:
Quality: 649.50 Length: 148
Ratio: 4.883 Gaps: 2
Percent Similarity: 89.865 Percent Identity: 81.757

alignment_block:
US-10-031-904-8 x AAN91477 ..

Align seg 1/1 to: AAN91477 from: 1 to: 6951

```
 10 ProPheProSerArgArgPhe......ProGlyLeuLeuLeuAlaAlaLe 24
    |||   |||:::     |||         |||   |||||||||:::::
 70 CCGGCGCCCGGTCTCCCCTTCTGCTGCGGAGGATCCCTGCTGGCGGTTGT 119

 24 uValLeuLeu...LeuSerSerPheSerAspGlnCysAsnValProGluT 40
    :|||||||||   |||               |||||||||:::|||||||
120 GGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAGAAT 169

 40 rpLeuProPheAlaArgProThrAsnLeuThrAspAspPheGluPhePro 56
    ||||||||||||||||||||||||||||||||||:::|||||||||||||
170 GGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCCC 219

 57 IleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgPr 73
    ||||||||||||||||||||||||||||||||||||||||||||||||||
220 ATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGACC 269

 73 oPheSerIleIleCysLeuLysAsnSerValTrpThrSerAlaLysAspL 90
    ||||||||||||||||||||||||||||||||||:::|||||||||:
270 GTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACTGGTGCTAAGGACA 319

 90 ysCysLysArgLysSerCysArgAsnProProAspProValAsnGlyMet 106
    ::|||:::||||||||||||||||||||||||||||||||||||||||||
320 GGTGCAGACGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATG 369

107 AlaHisValIleLysAspIleGlnPheGlySerGlnIleLysTyrSerCy 123
    :::||||||||||||   ||||||||||||||||||||||||||||||||
370 GTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAAATTAAATATTCTTG 419

123 sProLysGlyTyrArgLeuIleGlySerSerSerAlaThrCysIleIleS 140
    |   ||||||||||||||||||||||||||||||||||||||||||||||
420 TACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCATCT 469

140 erGlyAsnThrValIleTrpAspAsnLysThrProValCysAsp 154
    |||||:::|||||||||||||||||:::|||||||:::|||||||
470 CAGGTGATACTGTCATTTGGGATAATGAAACACCTATTTGTGAC 513
```

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:AAQ11642

seq_documentation_block:
ID AAQ11642 standard; DNA; 6951 BP.
XX
AC AAQ11642;
XX
DT 25-JUN-1991 (first entry)
XX
DE Entire human complement type 1 receptor coding region.
XX
KW complement system; C3b/C4b receptor; CR1; allergic reaction;
KW immune response; clone lambda T109.1; ss.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	sig_peptide	28..147
FT		/*tag= a
FT		/note= "putative"
FT	mat_peptide	148..6144
FT		/*tag= b
FT		/product= C3b/C4b receptor

XX
PN WO9105047-A.

```
XX
PD   18-APR-1991.
XX
PF   25-SEP-1990;   90WO-US05454.
XX
PR   26-SEP-1989;   89US-0412745.
PR   26-SEP-1990;   90US-0912349.
XX
PA   (TCEL-) T CELL SCI INC.
PA   (UYJO ) JOHNS HOPKINS UNIVERSITY.
PA   (BRIG-) BRIGHAM AND WOMEN'S HOSPITAL.
XX
PI   Fearon DT,  Klickstein LB,  Wong WW,  Carson GR,  Hoh M,  Concino MF;
PI   Makrides SC,  Marsh HC;
XX
DR   WPI; 1991-132854/18.
DR   P-PSDB; AAR11810.
XX
PT   Human complement receptor type 1 gene, encoded proteins and
PT   fragments - for treatment of immune disorders, myocardial infarct,
PT   damage due to inflammmation and in treatment of thrombosis
XX
PS   Claim 1; Fig 1; 234pp; English.
XX
CC   This sequence is a composite of sequences isolated as lambda clones
CC   H10.3, T109.1, H3.1 and H7.1. The clones were present in the
CC   specifically primed lambda gt11 cDNA library (Lambda HH) which was
CC   prepared with cDNA synthesised from poly(A)+ RNA from DMSO induced
CC   HL-60 cells. The library was screened using probes CR1-1, CR-2 and
CC   CR1-4 (see Wong,WW et al.,1985, Proc. Nat. Acad. Sci. USA. 82:7711)
CC   and probe CR1-18 (corresponding to nucleotides 101-352 of this
CC   sequence). There are four direct, long homologous repeats of 450bp,
CC   each comprising 7 short consensus repeats. Nucleotides 28-1533 are
CC   also claimed separately.
CC   See also AAQ11643.
XX
SQ   Sequence 6951 BP; 1802 A; 1681 C; 1659 G; 1809 T; 0 other;


alignment_scores:
            Quality:  649.50              Length:    148
              Ratio:  4.883                 Gaps:      2
 Percent Similarity:  89.865   Percent Identity: 81.757

alignment_block:
 US-10-031-904-8 x AAQ11642   ..

 Align seg 1/1  to: AAQ11642  from: 1  to: 6951

      10 ProPheProSerArgArgPhe......ProGlyLeuLeuLeuAlaAlaLe 24
         |||   |||:::     |||       |||   |||||||||:::::
      70 CCGGCGCCCGGTCTCCCCTTCTGCTGCGGAGGATCCCTGCTGGCGGTTGT 119

      24 uValLeuLeu...LeuSerSerPheSerAspGlnCysAsnValProGluT 40
         :|||||||||   |||                |||||||||:::|||||||
     120 GGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAGAAT 169

      40 rpLeuProPheAlaArgProThrAsnLeuThrAspAspPheGluPhePro 56
         ||||||||||||||||||||||||||||||||:::|||||||||||||
     170 GGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCCC 219

      57 IleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgPr 73
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     220 ATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGACC 269

      73 oPheSerIleIleCysLeuLysAsnSerValTrpThrSerAlaLysAspL 90
         ||||||||||||||||||||||||||||||||||:::|||||||||:
     270 GTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACTGGTGCTAAGGACA 319

      90 ysCysLysArgLysSerCysArgAsnProProAspProValAsnGlyMet 106
         ::|||:::||||||||||||||||||||||||||||||||||||||||||
     320 GGTGCAGACGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATG 369

     107 AlaHisValIleLysAspIleGlnPheGlySerGlnIleLysTyrSerCy 123
         :::|||||||||||||   |||||||||||||||||||||||||||||||
     370 GTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAAATTAAATATTCTTG 419

     123 sProLysGlyTyrArgLeuIleGlySerSerSerAlaThrCysIleIleS 140
         |   ||||||||||||||||||||||||||||||||||||||||||||||
     420 TACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCATCT 469

     140 erGlyAsnThrValIleTrpAspAsnLysThrProValCysAsp 154
         |||||:::|||||||||||||||||||:::||||||:::||||||
     470 CAGGTGATACTGTCATTTGGGATAATGAAACACCTATTTGTGAC 513


seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:AAZ38150

seq_documentation_block:
ID   AAZ38150 standard; DNA; 6951 BP.
XX
AC   AAZ38150;
XX
DT   22-FEB-2000  (first entry)
XX
DE   Human C3b/C4b receptor (CR1) protein encoding DNA.
XX
KW   C3B/C4B receptor; CR1 protein; cell-surface protein; erythrocyte; human;
KW   complement regulatory activity; complement pathway enzyme; tissue damage;
KW   reperfusion injury; Arthus reaction; myocardial infarct; inflammation;
KW   heart condition; autoimmune disorder; diagnostic; ss.
XX
OS   Homo sapiens.
XX
PN   US5981481-A.
XX
PD   09-NOV-1999.
XX
PF   06-JUN-1995;   95US-0470652.
XX
PR   03-APR-1989;   89US-0332865.
PR   06-DEC-1974;   74US-0350238.
PR   24-FEB-1993;   93US-0026134.
PR   01-APR-1988;   88US-0176532.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
PA   (AVAN-) AVANT IMMUNOTHERAPEUTICS INC.
XX
PI   Concino MF,  Wong WW,  Makrides SC,  Klickstein LB,  Fearon DT,  Ip SH;
PI   Marsh HC,  Carson GR;
XX
DR   WPI; 1999-633357/54.
DR   P-PSDB; AAY55751.
XX
PT   A human C3B/C4B receptor (CR1) protein having antiinflammatory and
PT   cardiant activity -
XX
PS   Disclosure; Fig 1A-P; 87pp; English.
XX
CC   The invention relates to a human C3B/C4B receptor (CR1) protein. The CR1
CC   protein or fragment is expressed as a cell-surface protein on the surface
CC   of a non-human cell and exhibits a complement regulatory activity of full
CC   -length human CR1 as expressed on erythrocytes. The CR1 function in vivo
CC   may be mediated through the inhibition of complement pathway enzymes. The
CC   soluble CR1 protein exhibits a complement regulatory activity, and this
CC   may be used to prevent reperfusion injury, inhibit Arthus reaction, and
CC   neutrophil mediated tissue damage, and reduce myocardial infarct size,
CC   and inflammation. The CR1 protein and its fragments can also be used in
CC   the treatment of conditions which involve unwanted complement activity,
CC   e.g. shock lung, tissue damage due to burn, or ischemic heart conditions,
CC   and autoimmune disorders. CR1 proteins, analogues, derivatives, and anti
CC   -CR1 antibodies are used in assays, and diagnostics. The present sequence
CC   represents a DNA encoding the human CR1 protein.
XX
SQ   Sequence 6951 BP; 1802 A; 1680 C; 1661 G; 1808 T; 0 other;
```

```
alignment_scores:
            Quality:  649.50              Length:    148
              Ratio:  4.883                 Gaps:      2
 Percent Similarity:  89.865   Percent Identity: 81.757

alignment_block:
 US-10-031-904-8 x AAZ38150   ..

 Align seg 1/1  to: AAZ38150  from: 1  to: 6951

      10 ProPheProSerArgArgPhe......ProGlyLeuLeuLeuAlaAlaLe 24
         |||   |||:::     |||         |||   |||||||||:::::
      70 CCGGCGCCCGGTCTCCCCTTCTGCTGCGGAGGATCCCTGCTGGCGGTTGT 119

      24 uValLeuLeu...LeuSerSerPheSerAspGlnCysAsnValProGluT 40
         :|||||||||   |||               ||||||||||:::|||||||
     120 GGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAGAAT 169

      40 rpLeuProPheAlaArgProThrAsnLeuThrAspAspPheGluPhePro 56
         ||||||||||||||||||||||||||||||||||:::|||||||||||||
     170 GGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCCC 219

      57 IleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgPr 73
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     220 ATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGACC 269

      73 oPheSerIleIleCysLeuLysAsnSerValTrpThrSerAlaLysAspL 90
         ||||||||||||||||||||||||||||||||||:::|||||||||:
     270 GTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACTGGTGCTAAGGACA 319

      90 ysCysLysArgLysSerCysArgAsnProProAspProValAsnGlyMet 106
         ::|||:::||||||||||||||||||||||||||||||||||||||||||
     320 GGTGCAGACGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATG 369

     107 AlaHisValIleLysAspIleGlnPheGlySerGlnIleLysTyrSerCy 123
         :::||||||||||||   ||||||||||||||||||||||||||||||||
     370 GTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAAATTAAATATTCTTG 419

     123 sProLysGlyTyrArgLeuIleGlySerSerSerAlaThrCysIleIleS 140
         |   ||||||||||||||||||||||||||||||||||||||||||||||||
     420 TACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCATCT 469

     140 erGlyAsnThrValIleTrpAspAsnLysThrProValCysAsp 154
         |||||:::|||||||||||||||||||:::|||||||:::||||||
     470 CAGGTGATACTGTCATTTGGGATAATGAAACACCTATTTGTGAC 513

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAI58380

seq_documentation_block:
ID   AAI58380 standard; cDNA; 7313 BP.
XX
AC   AAI58380;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 583.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM39224.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 583; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 7313 BP; 1903 A; 1770 C; 1733 G; 1907 T; 0 other;

alignment_scores:
            Quality:  649.50              Length:    148
              Ratio:  4.883                 Gaps:      2
 Percent Similarity:  89.865   Percent Identity: 81.757

alignment_block:
 US-10-031-904-8 x AAI58380   ..

 Align seg 1/1  to: AAI58380  from: 1  to: 7313

      10 ProPheProSerArgArgPhe......ProGlyLeuLeuLeuAlaAlaLe 24
         |||   |||:::     |||         |||   |||||||||:::::
      70 CCGGCGCCCGGTCTCCCCTTCTGCTGCGGAGGATCCCTGCTGGCGGTTGT 119

      24 uValLeuLeu...LeuSerSerPheSerAspGlnCysAsnValProGluT 40
         :|||||||||   |||               ||||||||||:::|||||||
     120 GGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAGAAT 169

      40 rpLeuProPheAlaArgProThrAsnLeuThrAspAspPheGluPhePro 56
         ||||||||||||||||||||||||||||||||||:::|||||||||||||
     170 GGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCCC 219

      57 IleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgPr 73
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     220 ATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGACC 269

      73 oPheSerIleIleCysLeuLysAsnSerValTrpThrSerAlaLysAspL 90
         ||||||||||||||||||||||||||||||||||:::|||||||||:
     270 GTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACTGGTGCTAAGGACA 319
```

```
 90 ysCysLysArgLysSerCysArgAsnProProAspProValAsnGlyMet 106
    ::|||:::||||||||||||||||||||||||||||||||||||||||||
320 GGTGCAGACGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATG 369

107 AlaHisValIleLysAspIleGlnPheGlySerGlnIleLysTyrSerCy 123
    :::|||||||||||||   ||||||||||||||||||||||||||||||
370 GTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAAATTAAATATTCTTG 419

123 sProLysGlyTyrArgLeuIleGlySerSerSerAlaThrCysIleIleS 140
    |   ||||||||||||||||||||||||||||||||||||||||||||||
420 TACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCATCT 469

140 erGlyAsnThrValIleTrpAspAsnLysThrProValCysAsp 154
    |||||:::|||||||||||||||||||:::||||||:::|||||
470 CAGGTGATACTGTCATTTGGGATAATGAAACACCTATTTGTGAC 513

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:ABA09026

seq_documentation_block:
ID   ABA09026 standard; cDNA; 7028 BP.
XX
AC   ABA09026;
XX
DT   11-JAN-2002  (first entry)
XX
DE   Human CR1 protein homologue-encoding cDNA, SEQ ID NO:802.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW   haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW   inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW   proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW   myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW   chronic inflammatory condition; proliferative retinopathy;
KW   atherosclerosis; coronary heart disease; arterial ischaemia;
KW   bone disorder; osteoporosis; vascular growth disorder;
KW   tissue regeneration; wound healing; infection; immune disorder;
KW   cell culture; drug screening; gene therapy; antiinflammatory;
KW   antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW   cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW   antifungal; vulnerary; antiulcer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157188-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US03800.
XX
PR   03-FEB-2000; 2000US-0496914.
PR   27-APR-2000; 2000US-0560875.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-457740/49.
DR   P-PSDB; ABB11782.
XX
PT   Human proteins and DNA encoding sequences useful for preventing,
PT   treating or ameliorating a medical condition in a mammalian subject
PT   e.g. arthritis and cancer -
XX
PS   Claim 1; Page 707-709; 1963pp; English.
XX
CC   Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC   sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC   invention also relates to vectors and recombinant host cells comprising a
CC   nucleotide of the invention, methods of producing the novel polypeptides,
CC   antibodies against the polypeptides, methods of detecting the nucleotides
CC   or polypeptides in a sample, and methods of identifying compounds which
CC   bind to polypeptides of the invention. Although novel, many of the
CC   polypeptides of the invention have homology to known proteins, thereby
CC   giving an insight into their probable biological activities, and hence
CC   potential therapeutic applications. The polypeptides of the invention may
CC   have various activities, including cytokine, cell proliferation or cell
CC   differentiation activities; stem cell growth factor activity;
CC   haematopoiesis regulatory activity; tissue growth activity;
CC   immunomodulatory activity; activin- or inhibin-related activities;
CC   chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC   thrombolytic activities; receptor or ligand activities; or may be
CC   involved in oncogenesis, cancer cell proliferation or metastasis.
CC   Depending on their biological activities, polypeptides and nucleotides of
CC   the invention are useful for preventing, treating or ameliorating medical
CC   conditions, e.g., by protein or gene therapy. Such conditions include
CC   cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC   disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC   proliferative retinopathy, atherosclerosis, coronary heart disease,
CC   arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC   vascular growth. Polypeptides involved with tissue regeneration and
CC   repair (or nucleic acids encoding them) may be used to promote wound
CC   healing (e.g., of burns, incisions and ulcers), while those with
CC   immunomodulatory activities may be used in the treatment of viral,
CC   bacterial and fungal infections in addition to immune disorders.
CC   Polypeptides with growth factor activity may be used in cell cultures to
CC   promote cell growth. For example, such polypeptides may be used to
CC   manipulate stem cells in culture to give rise to neuroepithelial cells
CC   that can be used to augment or replace cells damaged by illness,
CC   autoimmune disease or accidental damage. The polypeptides and nucleotides
CC   may also be used in the diagnosis of the above conditions, and in drug
CC   screening techniques. The present sequence represents a cDNA encoding a
CC   novel human polypeptide of the invention.
XX
SQ   Sequence 7028 BP; 1819 A; 1704 C; 1681 G; 1824 T; 0 other;

alignment_scores:
            Quality:  636.50              Length:    149
              Ratio:  4.786                 Gaps:      3
 Percent Similarity:  89.262   Percent Identity: 81.208

alignment_block:
 US-10-031-904-8 x ABA09026   ..

 Align seg 1/1  to: ABA09026  from: 1  to: 7028

  10 ProPheProSerArgArgPhe......ProGlyLeuLeuLeuAlaAlaLe 24
     |||   |||:::      |||         |||   ||||||||||:::::
  70 CCGGCGCCCGGTCTCCCCTTCTGCTGCGGAGGATCCCTGCTGGCGGTTGT 119

  24 uValLeuLeu...LeuSerSerPheSerAspGlnCysAsnValProGluT 40
     :|||||||||   |||                |||||||||:::|||||||
 120 GGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAGAAT 169

  40 rp.LeuProPheAlaArgProThrAsnLeuThrAspAspPheGluPhePr 56
     || ||||||||||||||||||||||||||||||||||::|||||||||||
 170 GGGCTTCCATTGCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCC 219

  56 oIleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgP 73
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 220 CATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGAC 269

  73 roPheSerIleIleCysLeuLysAsnSerValTrpThrSerAlaLysAsp 89
     ||||||||||||||||||||||||||||||||||:::|||||||||
 270 CGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACTGGTGCTAAGGAC 319

  90 LysCysLysArgLysSerCysArgAsnProProAspProValAsnGlyMe 106
     :::|||:::||||||||||||||||||||||||||||||||||||||||||
 320 AGGTGCAGACGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCAT 369

 106 tAlaHisValIleLysAspIleGlnPheGlySerGlnIleLysTyrSerC 123
     |:::|||||||||||||   ||||||||||||||||||||||||||||||
 370 GGTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAAATTAAATATTCTT 419

 123 ysProLysGlyTyrArgLeuIleGlySerSerSerAlaThrCysIleIle 139
     ||   ||||||||||||||||||||||||||||||||||||||||||||||
```

```
     420 GTACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCATC 469

     140 SerGlyAsnThrValIleTrpAspAsnLysThrProValCysAsp 154
         ||||||:::||||||||||||||||||:::||||||:::||||||
     470 TCAGGTGATACTGTCATTTGGGATAATGAAACACCTATTTGTGAC 514

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAI60166

seq_documentation_block:
ID   AAI60166 standard; cDNA; 7028 BP.
XX
AC   AAI60166;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 4155.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM41010.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 4155; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 7028 BP; 1819 A; 1704 C; 1681 G; 1824 T; 0 other;

alignment_scores:
            Quality:  636.50              Length:    149
              Ratio:  4.786                 Gaps:      3
 Percent Similarity:  89.262   Percent Identity: 81.208

alignment_block:
 US-10-031-904-8 x AAI60166   ..

 Align seq 1/1  to: AAI60166  from: 1  to: 7028

      10 ProPheProSerArgArgPhe......ProGlyLeuLeuLeuAlaAlaLe 24
         |||   |||:::      |||         |||   |||||||||:::::
      70 CCGGCGCCCGGTCTCCCCTTCTGCTGCGGAGGATCCCTGCTGGCGGTTGT 119

      24 uValLeuLeu...LeuSerSerPheSerAspGlnCysAsnValProGluT 40
         :|||||||||   |||               ||||||||||:::|||||||
     120 GGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAGAAT 169

      40 rp.LeuProPheAlaArgProThrAsnLeuThrAspAspPheGluPhePr 56
         || ||||||||||||||||||||||||||||||:::||||||||||||
     170 GGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCC 219

      56 oIleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgP 73
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     220 CATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGAC 269

      73 roPheSerIleIleCysLeuLysAsnSerValTrpThrSerAlaLysAsp 89
         ||||||||||||||||||||||||||||||||||||:::||||||||||
     270 CGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACTGGTGCTAAGGAC 319

      90 LysCysLysArgLysSerCysArgAsnProProAspProValAsnGlyMe 106
         :::|||:::||||||||||||||||||||||||||||||||||||||||
     320 AGGTGCAGACGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCAT 369

     106 tAlaHisValIleLysAspIleGlnPheGlySerGlnIleLysTyrSerC 123
         |:::|||||||||||||   ||||||||||||||||||||||||||||||
     370 GGTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAAATTAAATATTCTT 419

     123 ysProLysGlyTyrArgLeuIleGlySerSerSerAlaThrCysIleIle 139
         ||   ||||||||||||||||||||||||||||||||||||||||||||||
     420 GTACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCATC 469

     140 SerGlyAsnThrValIleTrpAspAsnLysThrProValCysAsp 154
         ||||||:::||||||||||||||||||:::||||||:::||||||
     470 TCAGGTGATACTGTCATTTGGGATAATGAAACACCTATTTGTGAC 514

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:AAV53262

seq_documentation_block:
ID   AAV53262 standard; DNA; 591 BP.
XX
AC   AAV53262;
XX
DT   18-JAN-1999  (first entry)
XX
DE   Complement receptor type 1-like sequence CM7 DNA.
XX
KW   Complement receptor type-1; CR1; CM7; complement; inhibitor;
KW   anti-haemolytic; multiple sclerosis; Parkinson's disease;
KW   xenograft rejection; inflammation; Crohn's disease; asthma;
KW   pancreatitis; post-ischaemic reperfusion; infection; sepsis;
KW   autoimmune disease; rheumatoid arthritis; proliferative nephritis;
KW   myasthenia gravis; reproductive disorder; therapy; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9839433-A1.
XX
PD   11-SEP-1998.
XX
```

PF 05-MAR-1998; 98WO-GB00727.
XX
PR 05-MAR-1997; 97GB-0004519.
XX
PA (ADPR-) ADPROTECH PLC.
XX
PI Cox VF, Mossakowska DEI, Smith RAG;
XX
DR WPI; 1998-506358/43.
DR P-PSDB; AAW79236.
XX
PT Soluble polypeptide comprising short consensus repeats from LHR-A -
PT used to treat disorders and diseases associated with inflammation or
PT inappropriate complement activation
XX
PS Claim 22; Page 42-43; 67pp; English.
XX
CC This DNA sequence encodes CM7 (see AAW79236), a protein that consists
CC of the short consensus repeats (SCR) 1 and 2 from the complement
CC receptor type 1 (CR1) fused to the SCR3 of CR1-like pseudogene (see
CC AAW79247). CM7 DNA was constructed using plasmid pDB1013-5, which
CC codes for SCR1-3 of CR1, by site-directed mutagenesis using 3 pairs
CC of oligonucleotides (see AAV53263-65) that introduced 10 amino acid
CC changes to the native SCR3 sequence corresponding to changes
CC observed in the CR1-like pseudogene (Cr1pse). pBrocSCR1-3CM7
CC carrying the CM7 DNA construct was used to transform Escherichia
CC coli BL21(DE3), and CM7 was purified from solubilised inclusion
CC bodies. The invention provides DNA sequences (see AAV53262 and
CC AAV53269-79) encoding novel soluble engineered CR1 polypeptides (see
CC AAW53236-47) such as CM7 that act as complement inhibitors with
CC functional complement inhibitory, including anti-haemolytic,
CC activity. These can be used to treat a disease or disorder
CC associated with inflammation or inappropriate complement activation,
CC such as neurological disorders (e.g. multiple sclerosis and
CC Parkinson's disease), disorders of inappropriate or undesirable
CC complement activation (e.g. xenograft rejection), inflammatory
CC disorders (e.g. Crohn's disease, asthma, and acute pancreatitis),
CC post-ischaemic reperfusion conditions, infection or sepsis,
CC immune complex disorders and autoimmune diseases (e.g. rheumatoid
CC arthritis, proliferative nephritis and myasthenia gravis), and
CC reproductive disorders.
XX
SQ Sequence 591 BP; 132 A; 159 C; 148 G; 152 T; 0 other;

alignment_scores:
Quality: 635.00 Length: 121
Ratio: 5.336 Gaps: 0
Percent Similarity: 98.347 Percent Identity: 90.909

alignment_block:
US-10-031-904-8 x AAV53262 ..

Align seg 1/1 to: AAV53262 from: 1 to: 591

```
      34 GlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuTh 50
         ||||||||||:::||||||||||||||||||||||||||||||||||||||
       4 CAGTGCAACGCTCCGGAATGGCTGCCGTTCGCGCGCCCGACCAACCTGAC 53

      50 rAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgP 67
         ||||:::||||||||||||||||||||||||||||||||||||||||||||
      54 TGATGAATTTGAGTTCCCGATCGGTACCTACCTGAACTACGAATGCCGCC 103

      67 roGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVal 83
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     104 CGGGTTATAGCGGCCGCCCGTTTTCTATCATCTGCCTGAAAAACTCTGTC 153

      84 TrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProPr 100
         ||||||:::||||||||:::|||:::||||||||||||||||||||||||
     154 TGGACTGGTGCTAAGGACCGTTGCCGACGTAAATCTTGTCGTAATCCGCC 203

      100 oAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlyS 117
          ||||||||||||||||||:::|||||||||||||   ||||||||||||
      204 AGATCCGGTTAACGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGTT 253

      117 erGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSer 133
          ||||||||||||||||||||   |||||||||||||||||||||||||||
      254 CCCAAATTAAATATTCTTGTACTAAAGGTTACCGTCTGATTGGTTCCTCC 303

      134 SerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysTh 150
          ||||||||||||||||||||||:::|||||||||||||||||||:::||
      304 AGCGCTACATGCATCATCTCTGGTGATACTGTCATTTGGGATAATGAAAC 353

      150 rProValCysAsp 154
          ||||:::||||||
      354 ACCGATTTGTGAC 366
```

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:AAV53269

seq_documentation_block:
ID AAV53269 standard; DNA; 591 BP.
XX
AC AAV53269;
XX
DT 18-JAN-1999 (first entry)
XX
DE Complement receptor type 1-like sequence CM1 DNA.
XX
KW Complement receptor type-1; CR1; CM1; complement; inhibitor;
KW anti-haemolytic; multiple sclerosis; Parkinson's disease;
KW xenograft rejection; inflammation; Crohn's disease; asthma;
KW pancreatitis; post-ischaemic reperfusion; infection; sepsis;
KW autoimmune disease; rheumatoid arthritis; proliferative nephritis;
KW myasthenia gravis; reproductive disorder; therapy; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO9839433-A1.
XX
PD 11-SEP-1998.
XX
PF 05-MAR-1998; 98WO-GB00727.
XX
PR 05-MAR-1997; 97GB-0004519.
XX
PA (ADPR-) ADPROTECH PLC.
XX
PI Cox VF, Mossakowska DEI, Smith RAG;
XX
DR WPI; 1998-506358/43.
DR P-PSDB; AAW79237.
XX
PT Soluble polypeptide comprising short consensus repeats from LHR-A -
PT used to treat disorders and diseases associated with inflammation or
PT inappropriate complement activation
XX
PS Claim 22; Page 44; 67pp; English.
XX
CC This DNA sequence encodes CM1 (see AAW79237), a protein that consists
CC of the short consensus repeats (SCR) 1 and 2 from complement
CC receptor type 1 (CR1) fused to an SCR3 (see AAW79242) in which 5 amino
CC acids were altered to those found in the SCR3 of the CR1-like
CC pseudogene (Cr1pse) putative product. CM1 DNA was constructed by
CC site-directed mutagenesis (see AAV53263) of plasmid pDB1013-5, which
CC codes for SCR1-3 of CR1. pBrocSCR1-3CM1 carrying CM1 DNA was used
CC to transform Escherichia coli BL21(DE3), and CM1 was purified from
CC solubilised inclusion bodies. The invention provides DNA sequences
CC (see AAV53262 and AAV53269-79) encoding novel soluble engineered CR1
CC polypeptides (see AAW53236-47) such as CM1 that act as complement
CC inhibitors with functional complement inhibitory, including
CC anti-haemolytic, activity. These can be used to treat a disease or
CC disorder associated with inflammation or inappropriate complement
CC activation, such as neurological disorders (e.g. multiple sclerosis
CC and Parkinson's disease), disorders of inappropriate or undesirable
CC complement activation (e.g. xenograft rejection), inflammatory

```
CC   disorders (e.g. Crohn's disease, asthma, and acute pancreatitis),
CC   post-ischaemic reperfusion conditions, infection or sepsis,
CC   immune complex disorders and autoimmune diseases (e.g. rheumatoid
CC   arthritis, proliferative nephritis and myasthenia gravis), and
CC   reproductive disorders.
XX
SQ   Sequence 591 BP; 127 A; 159 C; 151 G; 154 T; 0 other;


alignment_scores:
           Quality:  635.00            Length:    121
             Ratio:  5.336               Gaps:      0
 Percent Similarity:  98.347   Percent Identity: 90.909

alignment_block:
 US-10-031-904-8 x AAV53269   ..

 Align seg 1/1  to: AAV53269  from: 1  to: 591

      34 GlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuTh 50
         |||||||||:::||||||||||||||||||||||||||||||||||||||
       4 CAGTGCAACGCTCCGGAATGGCTGCCGTTCGCGCGCCCGACCAACCTGAC 53

      50 rAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgP 67
         ||||:::||||||||||||||||||||||||||||||||||||||||||||
      54 TGATGAATTTGAGTTCCCGATCGGTACCTACCTGAACTACGAATGCCGCC 103

      67 roGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVal 83
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     104 CGGGTTATAGCGGCCGCCCGTTTTCTATCATCTGCCTGAAAAACTCTGTC 153

      84 TrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProPr 100
         ||||||:::||||||||||:::|||:::||||||||||||||||||||||
     154 TGGACTGGTGCTAAGGACCGTTGCCGACGTAAATCTTGTCGTAATCCGCC 203

     100 oAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlyS 117
         ||||||||||||||||||::::||||||||||||   ||||||||||||||
     204 AGATCCGGTTAACGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGTT 253

     117 erGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSer 133
         ||||||||||||||||||||   ||||||||||||||||||||||||||||
     254 CCCAAATTAAATATTCTTGTACTAAAGGTTACCGTCTGATTGGTTCCTCC 303

     134 SerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysTh 150
         |||||||||||||||||||||||||:::|||||||||||||||||:::||
     304 AGCGCTACATGCATCATCTCTGGTGATACTGTCATTTGGGATAATGAAAC 353

     150 rProValCysAsp 154
         ||||:::|||||||
     354 ACCGATTTGTGAC 366

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:AAV53270

seq_documentation_block:
ID   AAV53270 standard; DNA; 591 BP.
XX
AC   AAV53270;
XX
DT   18-JAN-1999  (first entry)
XX
DE   Complement receptor type 1-like sequence CM2 DNA.
XX
KW   Complement receptor type-1; CR1; CM2; complement; inhibitor;
KW   anti-haemolytic; multiple sclerosis; Parkinson's disease;
KW   xenograft rejection; inflammation; Crohn's disease; asthma;
KW   pancreatitis; post-ischaemic reperfusion; infection; sepsis;
KW   autoimmune disease; rheumatoid arthritis; proliferative nephritis;
KW   myasthenia gravis; reproductive disorder; therapy; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9839433-A1.
XX
PD   11-SEP-1998.
XX
PF   05-MAR-1998;   98WO-GB00727.
XX
PR   05-MAR-1997;   97GB-0004519.
XX
PA   (ADPR-) ADPROTECH PLC.
XX
PI   Cox VF,  Mossakowska DEI,  Smith RAG;
XX
DR   WPI; 1998-506358/43.
DR   P-PSDB; AAW79238.
XX
PT   Soluble polypeptide comprising short consensus repeats from LHR-A -
PT   used to treat disorders and diseases associated with inflammation or
PT   inappropriate complement activation
XX
PS   Claim 22; Page 45; 67pp; English.
XX
CC   This DNA sequence encodes CM2 (see AAW79238), a protein that consists
CC   of the short consensus repeats (SCR) 1 and 2 from complement
CC   receptor type 1 (CR1) fused to an SCR3 (see AAW79243) in which 4 amino
CC   acids were altered to those found in the SCR3 of the CR1-like
CC   pseudogene (Crlpse) putative product.  CM2 DNA was constructed by
CC   site-directed mutagenesis (see AAV53264) of plasmid pDB1013-5, which
CC   codes for SCR1-3 of CR1.  pBrocSCR1-3CM2 carrying CM2 DNA was used
CC   to transform Escherichia coli BL21(DE3), and CM2 was purified from
CC   solubilised inclusion bodies.  The invention provides DNA sequences
CC   (see AAV53262 and AAV53269-79) encoding novel soluble engineered CR1
CC   polypeptides (see AAW53236-47) such as CM2 that act as complement
CC   inhibitors with functional complement inhibitory, including
CC   anti-haemolytic, activity.  These can be used to treat a disease or
CC   disorder associated with inflammation or inappropriate complement
CC   activation, such as neurological disorders (e.g. multiple sclerosis
CC   and Parkinson's disease), disorders of inappropriate or undesirable
CC   complement activation (e.g. xenograft rejection), inflammatory
CC   disorders (e.g. Crohn's disease, asthma, and acute pancreatitis),
CC   post-ischaemic reperfusion conditions, infection or sepsis,
CC   immune complex disorders and autoimmune diseases (e.g. rheumatoid
CC   arthritis, proliferative nephritis and myasthenia gravis), and
CC   reproductive disorders.
XX
SQ   Sequence 591 BP; 134 A; 158 C; 146 G; 153 T; 0 other;


alignment_scores:
           Quality:  635.00            Length:    121
             Ratio:  5.336               Gaps:      0
 Percent Similarity:  98.347   Percent Identity: 90.909

alignment_block:
 US-10-031-904-8 x AAV53270   ..

 Align seg 1/1  to: AAV53270  from: 1  to: 591

      34 GlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuTh 50
         |||||||||:::||||||||||||||||||||||||||||||||||||||
       4 CAGTGCAACGCTCCGGAATGGCTGCCGTTCGCGCGCCCGACCAACCTGAC 53

      50 rAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgP 67
         ||||:::||||||||||||||||||||||||||||||||||||||||||||
      54 TGATGAATTTGAGTTCCCGATCGGTACCTACCTGAACTACGAATGCCGCC 103

      67 roGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVal 83
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     104 CGGGTTATAGCGGCCGCCCGTTTTCTATCATCTGCCTGAAAAACTCTGTC 153

      84 TrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProPr 100
         ||||||:::||||||||||:::|||:::||||||||||||||||||||||
     154 TGGACTGGTGCTAAGGACCGTTGCCGACGTAAATCTTGTCGTAATCCGCC 203

     100 oAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlyS 117
```

```
     204 AGATCCGGTTAACGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGTT 253

     117 erGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSer 133
     254 CCCAAATTAAATATTCTTGTACTAAAGGTTACCGTCTGATTGGTTCCTCC 303

     134 SerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysTh 150
     304 AGCGCTACATGCATCATCTCTGGTGATACTGTCATTTGGGATAATGAAAC 353

     150 rProValCysAsp 154
     354 ACCGATTTGTGAC 366

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:AAV53271

seq_documentation_block:
ID   AAV53271 standard; DNA; 591 BP.
XX
AC   AAV53271;
XX
DT   18-JAN-1999  (first entry)
XX
DE   Complement receptor type 1-like sequence CM3 DNA.
XX
KW   Complement receptor type-1; CR1; CM3; complement; inhibitor;
KW   anti-haemolytic; multiple sclerosis; Parkinson's disease;
KW   xenograft rejection; inflammation; Crohn's disease; asthma;
KW   pancreatitis; post-ischaemic reperfusion; infection; sepsis;
KW   autoimmune disease; rheumatoid arthritis; proliferative nephritis;
KW   myasthenia gravis; reproductive disorder; therapy; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9839433-A1.
XX
PD   11-SEP-1998.
XX
PF   05-MAR-1998;   98WO-GB00727.
XX
PR   05-MAR-1997;   97GB-0004519.
XX
PA   (ADPR-) ADPROTECH PLC.
XX
PI   Cox VF,  Mossakowska DEI,  Smith RAG;
XX
DR   WPI; 1998-506358/43.
DR   P-PSDB; AAW79239.
XX
PT   Soluble polypeptide comprising short consensus repeats from LHR-A -
PT   used to treat disorders and diseases associated with inflammation or
PT   inappropriate complement activation
XX
PS   Claim 22; Page 46; 67pp; English.
XX
CC   This DNA sequence encodes CM2 (see AAW79239), a protein that consists
CC   of the short consensus repeats (SCR) 1 and 2 from complement
CC   receptor type 1 (CR1) fused to an SCR3 (see AAW79244) in which 1 amino
CC   acid was altered to that found in the SCR3 of the CR1-like
CC   pseudogene (Crlpse) putative product.  CM3 DNA was constructed by
CC   site-directed mutagenesis (see AAV53265) of plasmid pDB1013-5, which
CC   codes for SCR1-3 of CR1.  pBrocSCR1-3CM3 carrying CM3 DNA was used
CC   to transform Escherichia coli BL21(DE3), and CM3 was purified from
CC   solubilised inclusion bodies.  The invention provides DNA sequences
CC   (see AAV53262 and AAV53269-79) encoding novel soluble engineered CR1
CC   polypeptides (see AAW53236-47) such as CM3 that act as complement
CC   inhibitors with functional complement inhibitory, including
CC   anti-haemolytic, activity.  These can be used to treat a disease or
CC   disorder associated with inflammation or inappropriate complement
CC   activation, such as neurological disorders (e.g. multiple sclerosis
CC   and Parkinson's disease), disorders of inappropriate or undesirable
CC   complement activation (e.g. xenograft rejection), inflammatory
CC   disorders (e.g. Crohn's disease, asthma, and acute pancreatitis),
CC   post-ischaemic reperfusion conditions, infection or sepsis,
CC   immune complex disorders and autoimmune diseases (e.g. rheumatoid
CC   arthritis, proliferative nephritis and myasthenia gravis), and
CC   reproductive disorders.
XX
SQ   Sequence 591 BP; 131 A; 160 C; 149 G; 151 T; 0 other;


alignment_scores:
            Quality:  635.00            Length:    121
              Ratio:  5.336               Gaps:      0
 Percent Similarity:  98.347   Percent Identity: 90.909

alignment_block:
 US-10-031-904-8 x AAV53271   ..

 Align seg 1/1  to: AAV53271  from: 1  to: 591

      34 GlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuTh 50
       4 CAGTGCAACGCTCCGGAATGGCTGCCGTTCGCGCGCCCGACCAACCTGAC 53

      50 rAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgP 67
      54 TGATGAATTTGAGTTCCCGATCGGTACCTACCTGAACTACGAATGCCGCC 103

      67 roGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVal 83
     104 CGGGTTATAGCGGCCGCCCGTTTTCTATCATCTGCCTGAAAAACTCTGTC 153

      84 TrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProPr 100
     154 TGGACTGGTGCTAAGGACCGTTGCCGACGTAAATCTTGTCGTAATCCGCC 203

     100 oAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlyS 117
     204 AGATCCGGTTAACGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGTT 253

     117 erGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSer 133
     254 CCCAAATTAAATATTCTTGTACTAAAGGTTACCGTCTGATTGGTTCCTCC 303

     134 SerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysTh 150
     304 AGCGCTACATGCATCATCTCTGGTGATACTGTCATTTGGGATAATGAAAC 353

     150 rProValCysAsp 154
     354 ACCGATTTGTGAC 366

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:AAV53272

seq_documentation_block:
ID   AAV53272 standard; DNA; 591 BP.
XX
AC   AAV53272;
XX
DT   18-JAN-1999  (first entry)
XX
DE   Complement receptor type 1-like sequence CM5 DNA.
XX
KW   Complement receptor type-1; CR1; CM5; complement; inhibitor;
KW   anti-haemolytic; multiple sclerosis; Parkinson's disease;
KW   xenograft rejection; inflammation; Crohn's disease; asthma;
KW   pancreatitis; post-ischaemic reperfusion; infection; sepsis;
KW   autoimmune disease; rheumatoid arthritis; proliferative nephritis;
KW   myasthenia gravis; reproductive disorder; therapy; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
```

```
PN   WO9839433-A1.
XX
PD   11-SEP-1998.
XX
PF   05-MAR-1998;   98WO-GB00727.
XX
PR   05-MAR-1997;   97GB-0004519.
XX
PA   (ADPR-) ADPROTECH PLC.
XX
PI   Cox VF,  Mossakowska DEI,  Smith RAG;
XX
DR   WPI; 1998-506358/43.
DR   P-PSDB; AAW79240.
XX
PT   Soluble polypeptide comprising short consensus repeats from LHR-A -
PT   used to treat disorders and diseases associated with inflammation or
PT   inappropriate complement activation
XX
PS   Claim 22; Page 47; 67pp; English.
XX
CC   This DNA sequence encodes CM5 (see AAW79240), a protein that consists
CC   of the short consensus repeats (SCR) 1 and 2 from complement
CC   receptor type 1 (CR1) fused to an SCR3 (see AAW79245) in which 5 amino
CC   acids were altered to those found in the SCR3 of the CR1-like
CC   pseudogene (Crlpse) putative product.  CM5 DNA was constructed by
CC   site-directed mutagenesis (see AAV53264-65) of pDB1013-5, which
CC   codes for SCR1-3 of CR1.  pBrocSCR1-3CM5 carrying CM5 DNA was used
CC   to transform Escherichia coli BL21(DE3), and CM5 was purified from
CC   solubilised inclusion bodies.  The invention provides DNA sequences
CC   (see AAV53262 and AAV53269-79) encoding novel soluble engineered CR1
CC   polypeptides (see AAW53236-47) such as CM5 that act as complement
CC   inhibitors with functional complement inhibitory, including
CC   anti-haemolytic, activity.  These can be used to treat a disease or
CC   disorder associated with inflammation or inappropriate complement
CC   activation, such as neurological disorders (e.g. multiple sclerosis
CC   and Parkinson's disease), disorders of inappropriate or undesirable
CC   complement activation (e.g. xenograft rejection), inflammatory
CC   disorders (e.g. Crohn's disease, asthma, and acute pancreatitis),
CC   post-ischaemic reperfusion conditions, infection or sepsis,
CC   immune complex disorders and autoimmune diseases (e.g. rheumatoid
CC   arthritis, proliferative nephritis and myasthenia gravis), and
CC   reproductive disorders.
XX
SQ   Sequence 591 BP; 128 A; 160 C; 151 G; 152 T; 0 other;


alignment_scores:
            Quality:  635.00              Length:    121
              Ratio:  5.336                 Gaps:      0
 Percent Similarity:  98.347   Percent Identity: 90.909

alignment_block:
 US-10-031-904-8 x AAV53272   ..

 Align seg 1/1  to: AAV53272  from: 1  to: 591

      34 GlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuTh 50
         ||||||||||:::||||||||||||||||||||||||||||||||||||||
       4 CAGTGCAACGCTCCGGAATGGCTGCCGTTCGCGCGCCCGACCAACCTGAC 53

      50 rAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgP 67
         ||||:::||||||||||||||||||||||||||||||||||||||||||||
      54 TGATGAATTTGAGTTCCCGATCGGTACCTACCTGAACTACGAATGCCGCC 103

      67 roGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVal 83
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     104 CGGGTTATAGCGGCCGCCCGTTTTCTATCATCTGCCTGAAAAACTCTGTC 153

      84 TrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProPr 100
         ||||||:::||||||||||:::|||:::||||||||||||||||||||||||
     154 TGGACTGGTGCTAAGGACCGTTGCCGACGTAAATCTTGTCGTAATCCGCC 203

     100 oAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlyS 117
         ||||||||||||||||||:::||||||||||||   ||||||||||||
     204 AGATCCGGTTAACGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGTT 253

     117 erGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSer 133
         ||||||||||||||||||   ||||||||||||||||||||||||||||
     254 CCCAAATTAAATATTCTTGTACTAAAGGTTACCGTCTGATTGGTTCCTCC 303

     134 SerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysTh 150
         ||||||||||||||||||||||||:::|||||||||||||||||||:::||
     304 AGCGCTACATGCATCATCTCTGGTGATACTGTCATTTGGGATAATGAAAC 353

     150 rProValCysAsp 154
         ||||:::||||||
     354 ACCGATTTGTGAC 366

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:AAV53273

seq_documentation_block:
ID   AAV53273 standard; DNA; 591 BP.
XX
AC   AAV53273;
XX
DT   18-JAN-1999  (first entry)
XX
DE   Complement receptor type 1-like sequence CM6 DNA.
XX
KW   Complement receptor type-1; CR1; CM6; complement; inhibitor;
KW   anti-haemolytic; multiple sclerosis; Parkinson's disease;
KW   xenograft rejection; inflammation; Crohn's disease; asthma;
KW   pancreatitis; post-ischaemic reperfusion; infection; sepsis;
KW   autoimmune disease; rheumatoid arthritis; proliferative nephritis;
KW   myasthenia gravis; reproductive disorder; therapy; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9839433-A1.
XX
PD   11-SEP-1998.
XX
PF   05-MAR-1998;   98WO-GB00727.
XX
PR   05-MAR-1997;   97GB-0004519.
XX
PA   (ADPR-) ADPROTECH PLC.
XX
PI   Cox VF,  Mossakowska DEI,  Smith RAG;
XX
DR   WPI; 1998-506358/43.
DR   P-PSDB; AAW79241.
XX
PT   Soluble polypeptide comprising short consensus repeats from LHR-A -
PT   used to treat disorders and diseases associated with inflammation or
PT   inappropriate complement activation
XX
PS   Claim 22; Page 48; 67pp; English.
XX
CC   This DNA sequence encodes CM6 (see AAW79241), a protein that consists
CC   of the short consensus repeats (SCR) 1 and 2 from complement
CC   receptor type 1 (CR1) fused to an SCR3 (see AAW79246) in which 6 amino
CC   acids were altered to those found in the SCR3 of the CR1-like
CC   pseudogene (Crlpse) putative product.  CM6 DNA was constructed by
CC   site-directed mutagenesis (see AAV53263 and AAV53265) of pDB1013-5,
CC   which codes for SCR1-3 of CR1.  pBrocSCR1-3CM6 carrying CM6 DNA was used
CC   to transform Escherichia coli BL21(DE3), and CM6 was purified from
CC   solubilised inclusion bodies.  The invention provides DNA sequences
CC   (see AAV53262 and AAV53269-79) encoding novel soluble engineered CR1
CC   polypeptides (see AAW53236-47) such as CM6 that act as complement
CC   inhibitors with functional complement inhibitory, including
CC   anti-haemolytic, activity.  These can be used to treat a disease or
CC   disorder associated with inflammation or inappropriate complement
CC   activation, such as neurological disorders (e.g. multiple sclerosis
```

```
CC   and Parkinson's disease), disorders of inappropriate or undesirable
CC   complement activation (e.g. xenograft rejection), inflammatory
CC   disorders (e.g. Crohn's disease, asthma, and acute pancreatitis),
CC   post-ischaemic reperfusion conditions, infection or sepsis,
CC   immune complex disorders and autoimmune diseases (e.g. rheumatoid
CC   arthritis, proliferative nephritis and myasthenia gravis), and
CC   reproductive disorders.
XX
SQ   Sequence 591 BP; 135 A; 159 C; 146 G; 151 T; 0 other;


alignment_scores:
            Quality:  635.00            Length:    121
              Ratio:  5.336               Gaps:      0
 Percent Similarity:  98.347   Percent Identity: 90.909

alignment_block:
 US-10-031-904-8 x AAV53273   ..

 Align seg 1/1  to: AAV53273  from: 1  to: 591

      34 GlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuTh 50
         |||||||||:::|||||||||||||||||||||||||||||||||||||||
       4 CAGTGCAACGCTCCGGAATGGCTGCCGTTCGCGCGCCCGACCAACCTGAC 53

      50 rAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgP 67
         ||||:::|||||||||||||||||||||||||||||||||||||||||||
      54 TGATGAATTTGAGTTCCCGATCGGTACCTACCTGAACTACGAATGCCGCC 103

      67 roGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVal 83
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     104 CGGGTTATAGCGGCCGCCCGTTTTCTATCATCTGCCTGAAAAACTCTGTC 153

      84 TrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProPr 100
         ||||||:::|||||||||:::|||:::|||||||||||||||||||||||
     154 TGGACTGGTGCTAAGGACCGTTGCCGACGTAAATCTTGTCGTAATCCGCC 203

     100 oAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlyS 117
         ||||||||||||||||||:::|||||||||||   ||||||||||||
     204 AGATCCGGTTAACGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGTT 253

     117 erGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSer 133
         ||||||||||||||||||   |||||||||||||||||||||||||||||
     254 CCCAAATTAAATATTCTTGTACTAAAGGTTACCGTCTGATTGGTTCCTCC 303

     134 SerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysTh 150
         ||||||||||||||||||||||||:::|||||||||||||||||||:::||
     304 AGCGCTACATGCATCATCTCTGGTGATACTGTCATTTGGGATAATGAAAC 353

     150 rProValCysAsp 154
         ||||:::||||||
     354 ACCGATTTGTGAC 366

eq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:AAQ41867

eq_documentation_block:
ID   AAQ41867 standard; DNA; 6951 BP.
XX
AC   AAQ41867;
XX
DT   14-SEP-1993  (first entry)
XX
DE   CR1 coding region.
XX
KW   C3b/C4b receptor; CR1; erythrocyte; monocyte; macrophage; granulocyte;
KW   iC3b; T cell; splenic follicular dendritic cell; soluble; complement;
KW   glomerular podocyte; B cell; C3b; C4b; inactivated C3b; phagocytotsis;
KW   plasma; ligand binding activity; immune complex; activator; allotype;
KW   endocytosis; lymphocyte; classical; alternative; pathway; cofactor; F;
KW   C3/C5 convertase; liver; cleavage; factor I; regulation; glycoprotein;
KW   S; A; B; glycosylation; duplication; repetitive intervening sequence;
KW   endoglycosidase F; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             28..6147
FT                   /*tag= a
FT   sig_peptide     28..150
FT                   /*tag= b
FT   mat_peptide     151..6144
FT                   /*tag= c
FT   misc_RNA        1534..6147
FT                   /*tag= d
FT                   /note= "CR1 cDNA"
FT   misc_difference 509..512
FT                   /*tag= e
FT                   /note= "Unclear in the specification"
XX
PN   US5212071-A.
XX
PD   18-MAY-1993.
XX
PF   01-APR-1988;   88US-0176532.
XX
PR   01-APR-1988;   88US-0176532.
PR   03-APR-1989;   89US-0332865.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
PA   (TCEL-) T CELL SCI INC.
PA   (UYJO ) UNIV JOHNS HOPKINS.
XX
PI   Carson GR,  Concino MF,  Fearon DT,  Ip SH,  Klickstein LB;
PI   Makrides SC,  Wong WW;
XX
DR   WPI; 1993-175454/21.
DR   P-PSDB; AAR36743.
XX
PT   Nucleic acid encoding polypeptide having complement regulatory
PT   activity - used to prevent reperfusion injury, inhibit Arthus
PT   reaction and neutrophil mediated tissue damage and reduce
PT   myocardial infarct size and inflammation
XX
PS   Claim 1; Fig 1; 90pp; English.
XX
CC   This sequence represents the entire coding region for the C3b/C4b
CC   receptor (CR1).  CR1 is present on erythrocytes, monocytes/macro-
CC   phages, granulocytes, B cells, some T cells, splenic follicular
CC   dendritic cells and glomerular podocytes.  CR1 specifically binds
CC   C3b, C4b and inactivated C3b (iC3b).  A soluble form of the receptor
CC   is found in plasma which has ligand binding activity and the same
CC   molecular weight as membrane-associated CR1.  CR1 binds C3b and C4b
CC   that have covalently attached to immune complexes and other complement
CC   activators.  The consequences of these interactions depends on the
CC   type of bearing the receptor.  Erythrocyte CR1 binds immune complexes
CC   for transport to the liver.  CR1 on neutrophils and monocytes
CC   internalises bound complexes, either by adsorptive endocytosis
CC   or by phagocytotsis.  The function of CR1 on B lymphocytes is less
CC   well defined.  CR1 can inhibit the classical and alternative pathway
CC   C3/C5 convertases and act as a cofactor for the cleavage of C3b and
CC   C4b by factor I, therefore CR1 has a complement regulatory function
CC   as well as acting as a receptor.  CR1 is a glycoprotein composed of
CC   a single polypeptide chain.  Four allotypic forms of CR1 have been
CC   identified, differing by increments of approx. 40-50 kD.  The two most
CC   common forms, the F and S allotypes, also termed A and B allotypes,
CC   have molecular weights of 250 and 290 kD respectively.  The two rarer
CC   forms have molecular weights of 210 and 290 kD.  These differences
CC   represent differences in the polypeptide chain of CR1, rather than
CC   glycosylation state because they are not abolished by treatment of
CC   purified receptor portein with endoglycosidase F.  The CR1 gene has
CC   been shown to have repetitive intervening sequences which may have
CC   been duplicated in the formation of the larger allotypes.
XX
SQ   Sequence 6951 BP; 1799 A; 1692 C; 1648 G; 1807 T; 5 other;
```

```
alignment_scores:
            Quality:  634.50              Length:    146
              Ratio:  4.844                 Gaps:      2
 Percent Similarity:  89.726    Percent Identity: 81.507

alignment_block:
 US-10-031-904-8 x AAQ41867   ..

 Align seg 1/1  to: AAQ41867  from: 1  to: 6951

      10 ProPheProSerArgArgPhe.....ProGlyLeuLeuLeuAlaAlaLe 24
         |||   |||:::     |||        |||   |||||||||:::::
      70 CCGGCGCCCGGTCTCCCCTTCTGCTGCGGAGGATCCCTGCTGGCGGTTGT 119

      24 uValLeuLeu...LeuSerSerPheSerAspGlnCysAsnValProGluT 40
         :|||||||||   |||                |||||||||:::|||||||
     120 GGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAGAAT 169

      40 rpLeuProPheAlaArgProThrAsnLeuThrAspAspPheGluPhePro 56
         ||||||||||||||||||||||||||||||||:::|||||||||||||
     170 GGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCCC 219

      57 IleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgPr 73
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     220 ATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGACC 269

      73 oPheSerIleIleCysLeuLysAsnSerValTrpThrSerAlaLysAspL 90
         ||||||||||||||||||||||||||||||||||:::|||||||||:
     270 GTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACTGGTGCTAAGGACA 319

      90 ysCysLysArgLysSerCysArgAsnProProAspProValAsnGlyMet 106
         ::|||:::||||||||||||||||||||||||||||||||||||||||||
     320 GGTGCAGACGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATG 369

     107 AlaHisValIleLysAspIleGlnPheGlySerGlnIleLysTyrSerCy 123
         :::|||||||||||||   ||||||||||||||||||||||||||||||
     370 GTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAAATTAAATATTCTTG 419

     123 sProLysGlyTyrArgLeuIleGlySerSerSerAlaThrCysIleIleS 140
         |   ||||||||||||||||||||||||||||||||||||||||||||||
     420 TACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCATCT 469

     140 erGlyAsnThrValIleTrpAspAsnLysThrProVal 152
         |||||:::|||||||||||||||||:::|||||||:::
     470 CAGGTGATACTGTCATTTGGGATAATGAAACACCTATT 507
```

```
OM of: US-10-031-904-8 to: GenEmbl:*   out_format : pfs

Date: Oct 9, 2002  6:37 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlh
 -Q=/cgn2_1/USPTO_spool/US10031904/runat_09102002_084256_23638/app_query.fasta_1.232
 -DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 -USER=US10031904_@CGN1_1_5867 -NCPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-10-031-904-8
Query length: 174
Database: GenEmbl:*
Database sequences: 1797656
Database length: 1873333701
Search time (sec): 1840.610000


score_list:
Sequence           Strd Orig    ZScore    EScore Len  ! Documentation  ..
gb_pat:AX078362      +  946.00 1760.21  1.1e-89   627  ! AX078362 Sequence 30 from Paten
gb_pr:HSCR1          +  649.50 1182.47  1.7e-57  6951  ! Y00816 Human mRNA for compleme
gb_pr:CHPCR1X        +  645.50 1186.59  9.9e-58  1985  ! L24921 Pan troglodytes alterna
gb_pr:CHPCR1WT       +  645.50 1176.27  3.7e-57  6044  ! L24920 Chimpanzee complement r
gb_pr:HSCR1RS        +  644.50 1183.05  1.6e-57  2376  ! X14362 Human CR1 mRNA for C3b/
gb_sts:G28591        +  644.50 1183.05  1.6e-57  2376  ! G28591 human STS SHGC-35372, s
gb_pat:A86593        +  635.00 1178.14  2.9e-57   591  ! A86593 Sequence 2 from Patent W
gb_pat:A86601        +  635.00 1178.14  2.9e-57   591  ! A86601 Sequence 10 from Patent
gb_pat:A86603        +  635.00 1178.14  2.9e-57   591  ! A86603 Sequence 12 from Patent
gb_pat:A86605        +  635.00 1178.14  2.9e-57   591  ! A86605 Sequence 14 from Patent
gb_pat:A86607        +  635.00 1178.14  2.9e-57   591  ! A86607 Sequence 16 from Patent
gb_pat:A86609        +  635.00 1178.14  2.9e-57   591  ! A86609 Sequence 18 from Patent
gb_pat:AR029199      +  635.00 1177.93  3.0e-57   605  ! AR029199 Sequence 33 from paten
gb_pr:BABCORE        +  599.50 1101.92  5.2e-53  1688  ! L77977 Papio cynocephalus comp
gb_pr:BABCR1A        +  575.50 1045.21  7.4e-50  6000  ! L39791 Papio hamadryas complem
gb_pr:BABCOREA       +  561.00 1035.17  2.7e-49   945  ! L77978 Papio cynocephalus compl
gb_pr:CHPCR1Y        +  556.00 1020.19  1.8e-48  1731  ! L24922 Pan troglodytes mRNA se
gb_sts:G25967        -  551.50 1025.59  9.2e-49   389  ! G25967 human STS EST126198, seq
gb_sts:G27827        -  551.50 1025.59  9.2e-49   389  ! G27827 human STS SHGC-33387, se
gb_pr:HUMMCPLA       +  480.50  894.72  1.8e-41   309  ! M73721 Human MCP-like genetic e
gb_pat:AX333690      +  442.00  796.06  5.6e-36  5420  ! AX333690 Sequence 4199 from Pa
gb_pr:HSCR1R         +  442.00  796.06  5.6e-36  5420  ! X05309 Human mRNA for C3b/C4b
gb_ro:D42115         +  417.50  761.01  5.0e-34  1680  ! D42115 Rattus norvegicus mRNA
gb_ro:RAT5I2         +  417.50  760.77  5.2e-34  1724  ! D42114 Rat 5I2 antigen mRNA (c
gb_ro:RATCRRY        +  417.50  760.32  5.5e-34  1811  ! L36532 Rat complement regulato
gb_ro:MUSCRRYPS      +  398.00  727.06  3.9e-32  1272  ! M34174 Mouse complement recept
gb_pr:AL137789       +  395.50  678.48  2.0e-29 145325  ! AL137789 Human DNA sequence
gb_ro:MUSCRYL        +  394.50  717.38  1.4e-31  1781  ! M23529 Mus musculus complement
gb_ro:MUSXYG         +  393.00  720.01  9.6e-32   990  ! M16179 Mouse X/Y protein mRNA,
gb_htg:AL160003      +  388.00  656.88  3.2e-28 328209  ! AL160003 Homo sapiens chromo
gb_ro:GPICRRP        +  382.00  695.39  2.3e-30  1527  ! M77760 Guinea pig CRRP protein
gb_ro:MUSCR2AB       +  365.00  660.58  2.0e-28  2102  ! M36470 Mouse complement recept
gb_pr:HUMCR1L2       +  359.00  669.96  5.9e-29   227  ! M31231 Human complement recepto
gb_pr:HUMCR1SF02     +  342.00  627.75  1.3e-26   695  ! L17399 Human complement recepto
gb_pr:HUMCR102       +  334.00  623.13  2.4e-26   227  ! M31239 Human complement recepto
gb_ro:MUSCRYSPLN     +  293.50  540.54  9.6e-22   469  ! M23446 Mus musculus complement
gb_pat:AR146560      +  291.00  524.81  7.2e-21  1545  ! AR146560 Sequence 1 from paten
gb_pr:U87920         +  290.00  528.42  4.5e-21   855  ! U87920 Cercopithecus aethiops m
gb_pr:D63811         +  290.00  524.85  7.2e-21  1257  ! D63811 African green monkey mR
gb_pr:AF025483       +  289.00  530.55  3.4e-21   555  ! AF025483 Saimiri sciureus CD46
gb_pat:AR066585      +  289.00  526.04  6.2e-21   903  ! AR066585 Sequence 11 from pa
gb_pat:I70308        +  289.00  526.04  6.2e-21   903  ! I70308 Sequence 11 from pate
gb_pr:HUMSCD46       +  289.00  523.95  8.0e-21  1131  ! D84105 Human CD46 mRNA, com
gb_pat:AR112400      +  289.00  523.93  8.1e-21  1134  ! AR112400 Sequence 1 from pa
gb_pat:AR112402      +  289.00  523.93  8.1e-21  1134  ! AR112402 Sequence 3 from pa

seq_name: gb_pat:AX078362

seq_documentation_block:
LOCUS       AX078362               627 bp    DNA     linear   PAT 22-FEB-2001
DEFINITION  Sequence 30 from Patent WO0107612.
ACCESSION   AX078362
VERSION     AX078362.1  GI:13158031
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 627)
  AUTHORS   Au-Young,J., Bandman,O., Tang,Y.T., Yue,H., Azimzai,Y., Burford,N.,
            Baughn,M.R., Lu,D.A., Hillman,J.L., Patterson,C. and Lal,P.
  TITLE     Receptors and associated proteins
  JOURNAL   Patent: WO 0107612-A 30 01-FEB-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .627
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 103561CB1"
BASE COUNT      142 a    173 c    128 g    184 t
ORIGIN


alignment_scores:
            Quality:  946.00              Length:    174
              Ratio:  5.437                 Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-10-031-904-8 x AX078362   ..

 Align seg 1/1  to: AX078362  from: 1  to: 627

       1 MetAlaProProValArgLeuGluArgProPheProSerArgArgPhePr 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      87 ATGGCGCCTCCCGTCCGTCTCGAGCGTCCCTTTCCTTCCCGGCGCTTTCC 136

      17 oGlyLeuLeuLeuAlaAlaLeuValLeuLeuLeuSerSerPheSerAspG 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     137 TGGGTTGCTTCTGGCGGCCCTGGTGTTGCTGCTGTCCTCCTTCTCCGATC 186

      34 lnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThr 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     187 AATGCAATGTCCCGGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACT 236

      51 AspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgPr 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     237 GATGACTTTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCC 286

      67 oGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerValT 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     287 TGGTTATTCCGGAAGACCGTTTTCTATCATCTGCCTAAAAAACTCAGTCT 336

      84 rpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProPro 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     337 GGACAAGTGCTAAGGACAAGTGCAAACGTAAATCATGTCGTAATCCTCCA 386

     101 AspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlySe 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     387 GATCCTGTGAATGGCATGGCACATGTGATCAAAGACATCCAGTTCGGATC 436

     117 rGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSerS 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
     437 CCAAATTAAATATTCTTGTCCTAAAGGATACCGACTCATTGGTTCCTCGT 486

     134 erAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysThr 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     487 CTGCCACATGCATCATCTCAGGCAACACTGTCATTTGGGATAATAAAACA 536

     151 ProValCysAspSerGluLeuLysTyrAlaPheLeuPheLeuLeuProIl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     537 CCTGTTTGTGACAGTGAGTTGAAATATGCATTCCTATTTCTTTTACCGAT 586

     167 eHisSerAsnPheSerLeuGlu 174
         ||||||||||||||||||||||
     587 ACATTCTAATTTTTCTCTGGAA 608

seq_name: gb_pr:HSCR1

seq_documentation_block:
LOCUS       HSCR1                   6951 bp    mRNA    linear   PRI 23-MAR-1995
DEFINITION  Human mRNA for complement receptor type 1 (CR1, C3b/C4b receptor,
            CD35).
ACCESSION   Y00816
VERSION     Y00816.1  GI:30185
KEYWORDS    C3b/C4b complement component receptor; C3b/C4b receptor; CD35
            antigen; complement receptor; glycoprotein; membrane protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6951)
  AUTHORS   Klickstein,L.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-1988) Klickstein L.B., Room 617 Hunterian Bldg.,
            725 N. Wolfe St., Baltimore, MD 21205
REFERENCE   2  (bases 1 to 1531)
  AUTHORS   Klickstein,L.B., Bartow,T.J., Miletic,V., Rabson,L.D., Smith,J.A.
            and Fearon,D.T.
  TITLE     Identification of distinct C3b and C4b recognition sites in the
            human C3b/C4b receptor (CR1, CD35) by deletion mutagenesis
  JOURNAL   J. Exp. Med. 168 (5), 1699-1717 (1988)
  MEDLINE   89035992
COMMENT     This is the sequence of the F allotype of CR1.
            seq pos. 1532-6951 are already published see x05309 Data kindly
            reviewed (16/5/89) by Klickstein L.
FEATURES             Location/Qualifiers
     source          1. .6951
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /tissue_type="tonsil and HL-60"
                     /clone_lib="lambda S2T (ATCC #37546)."
     sig_peptide     28. .150
     CDS             28. .6147
                     /note="CR1 precursor protein"
                     /codon_start=1
                     /protein_id="CAA68755.1"
                     /db_xref="GI:30186"
                     /db_xref="SWISS-PROT:P17927"
                     /translation="MGASSPRSPEPVGPPAPGLPFCCGGSLLAVVVLLALPVAWGQCN
                     APEWLPFARPTNLTDEFEFPIGTYLNYECRPGYSGRPFSIICLKNSVWTGAKDRCRRK
                     SCRNPPDPVNGMVHVIKGIQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIWDNETPI
                     CDRIPCGLPPTITNGDFISTNRENFHYGSVVTYRCNPGSGGRKVFELVGEPSIYCTSN
                     DDQVGIWSGPAPQCIIPNKCTPPNVENGILVSDNRSLFSLNEVVEFRCQPGFVMKGPR
                     RVKCQALNKWEPELPSCSRVCQPPPDVLHAERTQRDKDNFSPGQEVFYSCEPGYDLRG
                     AASMRCTPQGDWSPAAPTCEVKSCDDFMGQLLNGRVLFPVNLQLGAKVDFVCDEGFQL
                     KGSSASYCVLAGMESLWNSSVPVCEQIFCPSPPVIPNGRHTGKPLEVFPFGKAVNYTC
                     DPHPDRGTSFDLIGESTIRCTSDPQGNGVWSSPAPRCGILGHCQAPDHFLFAKLKTQT
                     NASDFPIGTSLKYECRPEYYGRPFSITCLDNLVWSSPKDVCKRKSCKTPPDPVNGMVH
                     VITDIQVGSRINYSCTTGHRLIGHSSAECILSGNAAHWSTKPPICQRIPCGLPPTIAN
                     GDFISTNRENFHYGSVVTYRCNPGSGGRKVFELVGEPSIYCTSNDDQVGIWSGPAPQC
                     IIPNKCTPPNVENGILVSDNRSLFSLNEVVEFRCQPGFVMKGPRRVKCQALNKWEPEL
                     PSCSRVCQPPPDVLHAERTQRDKDNFSPGQEVFYSCEPGYDLRGAASMRCTPQGDWSP
                     AAPTCEVKSCDDFMGQLLNGRVLFPVNLQLGAKVDFVCDEGFQLKGSSASYCVLAGME
                     SLWNSSVPVCEQIFCPSPPVIPNGRHTGKPLEVFPFGKAVNYTCDPHPDRGTSFDLIG
                     ESTIRCTSDPQGNGVWSSPAPRCGILGHCQAPDHFLFAKLKTQTNASDFPIGTSLKYE
                     CRPEYYGRPFSITCLDNLVWSSPKDVCKRKSCKTPPDPVNGMVHVITDIQVGSRINYS
                     CTTGHRLIGHSSAECILSGNTAHWSTKPPICQRIPCGLPPTIANGDFISTNRENFHYG
                     SVVTYRCNLGSRGRKVFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNVENG
                     ILVSDNRSLFSLNEVVEFRCQPGFVMKGPRRVKCQALNKWEPELPSCSRVCQPPPEIL
                     HGEHTPSHQDNFSPGQEVFYSCEPGYDLRGAASLHCTPQGDWSPEAPRCAVKSCDDFL
                     GQLPHGRVLFPLNLQLGAKVSFVCDEGFRLKGSSVSHCVLVGMRSLWNNSVPVCEHIF
                     CPNPPAILNGRHTGTPSGDIPYGKEISYTCDPHPDRGMTFNLIGESTIRCTSDPHGNG
                     VWSSPAPRCELSVRAGHCKTPEQFPFASPTIPINDFEFPVGTSLNYECRPGYFGKMFS
                     ISCLENLVWSSVEDNCRRKSCGPPPEPFNGMVHINTDTQFGSTVNYSCNEGFRLIGSP
                     STTCLVSGNNVTWDKKAPICEIISCEPPPTISNGDFYSNNRTSFHNGTVVTYQCHTGP
                     DGEQLFELVGERSIYCTSKDDQVGVWSSPPPRCISTNKCTAPEVENAIRVPGNRSFFS
                     LTEIIRFRCQPGFVMVGSHTVQCQTNGRWGPKLPHCSRVCQPPPEILHGEHTLSHQDN
                     FSPGQEVFYSCEPSYDLRGAASLHCTPQGDWSPEAPRCTVKSCDDFLGQLPHGRVLLP
                     LNLQLGAKVSFVCDEGFRLKGRSASHCVLAGMKALWNSSVPVCEQIFCPNPPAILNGR
                     HTGTPFGDIPYGKEISYACDTHPDRGMTFNLIGESSIRCTSDPQGNGVWSSPAPRCEL
                     SVPAACPHPPKIQNGHYIGGHVSLYLPGMTISYTCDPGYLLVGKGFIFCTDQGIWSQL
                     DHYCKEVNCSFPLFMNGISKELEMKKVYHYGDYVTLKCEDGYTLEGSPWSQCQADDRW
                     DPPLAKCTSRAHDALIVGTLSGTIFFILLIIFLSWIILKHRKGNNAHENPKEVAIHLH
                     SQGGSSVHPRTLQTNEENSRVLP"
     mat_peptide     151. .6144
                     /product="mature CR1 protein (AA 1-1998)"
     misc_feature    151. .1500
                     /note="long homologous repeat A coding sequence"
     misc_feature    1501. .2850
                     /note="long homologous repeat B coding sequence"
     misc_feature    2851. .4209
                     /note="long homologous repeat C coding sequence"
     misc_feature    4210. .5565
                     /note="long homologous repeat D coding sequence"
BASE COUNT     1802 a   1680 c   1661 g   1808 t
ORIGIN


alignment_scores:
           Quality:  649.50            Length:    148
             Ratio:  4.883               Gaps:      2
 Percent Similarity:  89.865   Percent Identity: 81.757

alignment_block:
 US-10-031-904-8 x HSCR1   ..

 Align seg 1/1  to: HSCR1  from: 1  to: 6951

      10 ProPheProSerArgArgPhe......ProGlyLeuLeuLeuAlaAlaLe 24
         |||   |||:::     |||        |||   |||||||||:::::
      70 CCGGCGCCCGGTCTCCCCTTCTGCTGCGGAGGATCCCTGCTGGCGGTTGT 119

      24 uValLeuLeu...LeuSerSerPheSerAspGlnCysAsnValProGluT 40
         :|||||||||   |||               |||||||||:::|||||||
     120 GGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAGAAT 169

      40 rpLeuProPheAlaArgProThrAsnLeuThrAspAspPheGluPhePro 56
         ||||||||||||||||||||||||||||||:::||||||||||||||||
     170 GGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCCC 219

      57 IleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgPr 73
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     220 ATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGACC 269

      73 oPheSerIleIleCysLeuLysAsnSerValTrpThrSerAlaLysAspL 90
         ||||||||||||||||||||||||||||||||||:::|||||||||:
     270 GTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACTGGTGCTAAGGACA 319

      90 ysCysLysArgLysSerCysArgAsnProProAspProValAsnGlyMet 106
         ::|||:::||||||||||||||||||||||||||||||||||||||||||
     320 GGTGCAGACGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATG 369

     107 AlaHisValIleLysAspIleGlnPheGlySerGlnIleLysTyrSerCy 123
         :::|||||||||||   ||||||||||||||||||||||||||||||||
     370 GTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAAATTAAATATTCTTG 419

     123 sProLysGlyTyrArgLeuIleGlySerSerSerAlaThrCysIleIleS 140
         |   ||||||||||||||||||||||||||||||||||||||||||||||
```

```
    420 TACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCATCT 469

    140 erGlyAsnThrValIleTrpAspAsnLysThrProValCysAsp 154
        |||||:::|||||||||||||||||||:::||||||:::||||||
    470 CAGGTGATACTGTCATTTGGGATAATGAAACACCTATTTGTGAC 513

seq_name: gb_pr:CHPCR1X

seq_documentation_block:
LOCUS       CHPCR1X                 1985 bp    mRNA    linear   PRI 30-OCT-1994
DEFINITION  Pan troglodytes alternatively spliced CR1 (CR1) gene, partial cds.
ACCESSION   L24921
VERSION     L24921.1  GI:557726
KEYWORDS    alternative splicing product; complement receptor 1.
SOURCE      Pan troglodytes cDNA to mRNA.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 1985)
  AUTHORS   Birmingham,D.J., Shen,X.P., Hourcade,D., Nickells,M.W. and
            Atkinson,J.P.
  TITLE     Primary sequence of an alternatively spliced form of CR1. Candidate
            for the 75,000 M(r) complement receptor expressed on chimpanzee
            erythrocytes
  JOURNAL   J. Immunol. 153 (2), 691-700 (1994)
  MEDLINE   94292799
FEATURES             Location/Qualifiers
     source          1. .1985
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
                     /cell_line="EBV transformed"
                     /cell_type="B cell"
     sig_peptide     <1. .50
                     /gene="CR1"
     CDS             <1. .>1985
                     /gene="CR1"
                     /codon_start=3
                     /product="complement receptor 1"
                     /protein_id="AAA51439.1"
                     /db_xref="GI:557727"
                     /translation="SLLAVVVLLALPVAWGQCNAPEWLPFARPTNLTDEFEFPIGTYL
                     NYECRPGYYGRPFSIICLKNSVWTGAKDRCRRKSCRNPPDPVNGMVHVIKDIQFGSQI
                     KYSCTKGYRLIGSSSATCIISGDTVIWDNETPICDRIPCGLPPTITNGDFISTNRENF
                     HYGSVVTYRCNPGSGGRKVFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNV
                     ENGILVSDNRSLFSLNEVVEFRCQPGFVMKGPPRVKCQALNKWEPELPSCSRVCQPPP
                     DVLHAERTQRDKDNFSPGQEVFYSCEPGYDLRGAASLRCTPQGDWSPATPTCEVKSCD
                     DFMGQLLNGRVLFPVNLQLGAKVDFVCDEGFQLKGSSASYCVLAGMESLWNSSVPVCE
                     QIFCPNPPAILNGRHTGTPFGDIPYGKEISYACDTHPDRGMTFNLIGESSIRCTSDPQ
                     GNGVWSSPAPRCELSVPAACPHPPKIQNGHDIGGHVSLYLPGMTISYICDPGYLLVGK
                     GFIFCTDQGIWSQLDHYCKEVNCSFPLFMNGISKELEMKKVYHYGDYVTLKCEDGYTL
                     EGSPWSQCQADDRWDPPLAKCTSRAHDALIVGTSSGTIFFILLIIFLSWIILKHRKGN
                     NAHEKPKEVAIHLHSQGGSSVHPRTLQTNEENSRVLP"
     gene            1. .1985
                     /gene="CR1"
     mat_peptide     51. .1985
                     /gene="CR1"
                     /product="complement receptor 1"
BASE COUNT      521 a    476 c    478 g    510 t
ORIGIN


alignment_scores:
            Quality:  645.50             Length:    136
              Ratio:  5.043                Gaps:      1
 Percent Similarity:  94.118   Percent Identity: 86.029

alignment_block:
 US-10-031-904-8 x CHPCR1X   ..

 Align seg 1/1  to: CHPCR1X  from: 1  to: 1985

      20 LeuLeuAlaAlaLeuValLeuLeu...LeuSerSerPheSerAspGlnCy 35
         |||||||||:::::|||||||||   |||               |||||

      6 CTGCTGGCGGTTGTGGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATG 55

     35 sAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThrAspA 52
        ||||:::|||||||||||||||||||||||||||||||||||||||||:
     56 CAATGCCCCAGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATG 105

     52 spPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgProGly 68
        ::|||||||||||||||||||||||||||||||||||||||||||||||
    106 AGTTTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGT 155

     69 TyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerValTrpTh 85
        |||   |||||||||||||||||||||||||||||||||||||||||||
    156 TATTATGGAAGACCGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGAC 205

     85 rSerAlaLysAspLysCysLysArgLysSerCysArgAsnProProAspP 102
        |:::|||||||||:::|||:::|||||||||||||||||||||||||||
    206 TGGTGCTAAGGACAGGTGCAGACGTAAATCATGTCGTAATCCTCCAGATC 255

    102 roValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlySerGln 118
        ||||||||||||||:::|||||||||||||||||||||||||||||||
    256 CTGTGAATGGCATGGTGCATGTGATCAAAGACATCCAATTCGGATCCCAA 305

    119 IleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSerSerAl 135
        |||||||||||||||   ||||||||||||||||||||||||||||||
    306 ATTAAATATTCTTGTACTAAAGGATACCGACTCATTGGTTCCTCGTCTGC 355

    135 aThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysThrProV 152
        ||||||||||||||||||:::|||||||||||||||||:::||||||:
    356 CACATGCATCATCTCAGGTGATACTGTCATTTGGGATAATGAAACACCTA 405

    152 alCysAsp 154
        ::|||||
    406 TTTGTGAC 413

seq_name: gb_pr:CHPCR1WT

seq_documentation_block:
LOCUS       CHPCR1WT                6044 bp    mRNA    linear   PRI 30-OCT-1994
DEFINITION  Chimpanzee complement receptor type one (CR1) mRNA, partial cds.
ACCESSION   L24920
VERSION     L24920.1  GI:551564
KEYWORDS    complement receptor 1.
SOURCE      Pan troglodytes cDNA to mRNA.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 6044)
  AUTHORS   Birmingham,D.J., Shen,X.P., Hourcade,D., Nickells,M.W. and
            Atkinson,J.P.
  TITLE     Primary sequence of an alternatively spliced form of CR1. Candidate
            for the 75,000 M(r) complement receptor expressed on chimpanzee
            erythrocytes
  JOURNAL   J. Immunol. 153 (2), 691-700 (1994)
  MEDLINE   94292799
FEATURES             Location/Qualifiers
     source          1. .6044
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
     gene            1. .6044
                     /gene="CR1"
     CDS             <1. .>6044
                     /gene="CR1"
                     /codon_start=3
                     /product="complement receptor 1"
                     /protein_id="AAA51438.1"
                     /db_xref="GI:557725"
                     /translation="SLLAVVVLLALPVAWGQCNAPEWLPFARPTNLTDEFEFPIGTYL
                     NYECRPGYYGRPFSIICLKNSVWTGAKDRCRRKSCRNPPDPVNGMVHVIKDIQFGSQI
                     KYSCTKGYRLIGSSSATCIISGDTVIWDNETPICDRIPCGLPPTITNGDFISTNRENF
                     HYGSVVTYRCNPGSGGRKVFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNV
                     ENGILVSDNRSLFSLNEVVEFRCQPGFVMKGPPRVKCQALNKWEPELPSCSRVCQPPP
                     DVLHAERTQRDKDNFSPGQEVFYSCEPGYDLRGAASLRCTPQGDWSPATPTCEVKSCD
                     DFMGQLLNGRVLFPVNLQLGAKVDFVCDEGFQLKGSSASYCVLAGMESLWNSSVPVCE
```

QIFCPSPPVIPNGRHTGKPLEVFPFGKAVNYTCDPHPDRGTTFDLIGESTIRCTSDPQ
GNGVWSSPAPRCGILGHCQAPDHFLFAKLKTQTNASDFPIGTSLKYKCRPEYYGRPFS
ITCLDNLVWSSPKDVCKRKSCKTPPDPVNGMVHVITDIQVGSRINYSCTTGHRLIGHS
SAECILSGNAAHWSTKPPICQRIPCGLPPTIANGDFISTNRENFHYGSVVTYRCNPGS
RGRKVFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNVENGILVSDNRSLFS
LNEVVEFRCQPGFAMKGPRRVKCQALNKWEPELPSCSRVCQPPPDVLHAERTQRDKDN
FSPGQEVFYSCEPGYDLRGAASLRCTPQGDWSPAAPTCEVKSCDGFMGQLLNGRVLFP
VNLQLGAKVDFVCDEGFQLKGSSASYCVLAGMESLWNSSVPVCEQIFCPSPPVIPNGR
HTGKPLEVFPFGKAVNYTCDPHPDRGTTFDLIGESTIRCTSDPQGNGVWSSPAPRCGI
LGHCQAPDHFLFAKLKTQTNASDFPIGTSLKYECRPEYYGRPFSITCLDNLVWSSPKD
VCRRKSCKTPPDPVNGMVHVITDIQVGSRINYSCTTGHRLIGHSSAECILSGNSAHWS
TKPPICQRIPCGLPPTIANGDFISTNRENFHYGSVVTYRCNLGSRGRKVFELVGEPSI
YCTSNDDQVGIWSGPAPQCIIPNKCTPPNVENGILVSDNRSLFSLNEVVEFRCQPGFV
MKGPHRVKCQALNKWEPELPSCSRVCQPPPEILHGEHTPSHQDNFSPGQEVFYSCEPG
YDLRGAASLHCTPQGDWSPEAPRCAVKSCDDFLGQLPHGRVLFPLNLQLGAKVSFVCD
EGFRLKGSSVSHCVLVGMRSLWNNSVPVCEQIFCPNPPAILNGRHTGTPFGDIPYGKE
ISYTCDPHPDRGMTFNLIGESTIRCTSDPHGNGVWSSPAPRCELPVHAGHCKTPEQFP
FASPTIPINDFEFPVGTSLNYECRPGYFGKMFSISCLENLVWSSVEDNCRRKSCGPPP
EPFNGMVHINTDTQFGSTVNYSCNEGFRLIGSPSTTCLVSGNNVTWDKKAPICEIISC
EPPPTISNGDFYSNNRASFHNGTVVTYQCHTGPDGEQLFELVGERSIYCTSKDDQVGV
WSSPPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEIVRFRCQPGFVMVGSHTVQCQT
NGRWGPKLPHCSRVCQPPPEILHGEHTPSHQDNFSPGQEVFYSCEPGYDLRGAASLHC
TPQGDWSPEAPRCTVKSCDDFLGQLPHGRVLFPLNLQLGAKVSFVCDEGFRLKGRSAS
HCVLAGMKALWNSSVPVCEQIFCPNPPAILNGRHTGTPFGDIPYGKEISYACDTHPDR
GMTFNLIGESSIRCTSDPQGNGVWSSPAPRCELSVPAACPHPPKIQNGHDIGGHVSLY
LPGMTISYICDPGYLLVGKGFIFCTDQGIWSQLDHYCKEVNCSFPLFMNGISKELEMK
KVYHYGDYVTLKCEDGYTLEGSPWSQCQADDRWDPPLAKCTSRAHDALIVGTSSGTIF
FILLIIFLSWIILKHRKGNNAHEKPKEVAIHLHSQGGSSVHPRTLQTNEENSRVLP"

BASE COUNT 1561 a 1496 c 1462 g 1525 t
ORIGIN

alignment_scores:
Quality: 645.50 Length: 136
Ratio: 5.043 Gaps: 1
Percent Similarity: 94.118 Percent Identity: 86.029

alignment_block:
US-10-031-904-8 x CHPCR1WT ..

Align seg 1/1 to: CHPCR1WT from: 1 to: 6044

```
      20 LeuLeuAlaAlaLeuValLeuLeu...LeuSerSerPheSerAspGlnCy 35
         |||||||||:::::|||||||||   |||                |||||
       6 CTGCTGGCGGTTGTGGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATG 55

      35 sAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThrAspA 52
         ||||:::|||||||||||||||||||||||||||||||||||||||||||:
      56 CAATGCCCCAGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATG 105

      52 spPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgProGly 68
         ::|||||||||||||||||||||||||||||||||||||||||||||||||
     106 AGTTTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGT 155

      69 TyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerValTrpTh 85
         |||   |||||||||||||||||||||||||||||||||||||||||||||
     156 TATTATGGAAGACCGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGAC 205

      85 rSerAlaLysAspLysCysLysArgLysSerCysArgAsnProProAspP 102
         |:::|||||||||:::|||:::|||||||||||||||||||||||||||||
     206 TGGTGCTAAGGACAGGTGCAGACGTAAATCATGTCGTAATCCTCCAGATC 255

     102 roValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlySerGln 118
         ||||||||||||||:::||||||||||||||||||||||||||||||||||
     256 CTGTGAATGGCATGGTGCATGTGATCAAAGACATCCAATTCGGATCCCAA 305

     119 IleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSerSerAl 135
         ||||||||||||||   |||||||||||||||||||||||||||||||||
     306 ATTAAATATTCTTGTACTAAAGGATACCGACTCATTGGTTCCTCGTCTGC 355

     135 aThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysThrProV 152
         ||||||||||||||||||:::||||||||||||||||:::|||||||:
     356 CACATGCATCATCTCAGGTGATACTGTCATTTGGGATAATGAAACACCTA 405

      152 alCysAsp 154
          ::|||||
      406 TTTGTGAC 413
```

seq_name: gb_pr:HSCR1RS

seq_documentation_block:

LOCUS HSCR1RS 2376 bp mRNA linear PRI 22-MAR-1995
DEFINITION Human CR1 mRNA for C3b/C4b receptor secreted form.
ACCESSION X14362 Y00812
VERSION X14362.1 GI:30197
KEYWORDS alternate splicing; C3b/C4b receptor; complement receptor; receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2376)
AUTHORS Hourcade,D.
TITLE Direct Submission
JOURNAL Submitted (29-NOV-1988) Hourcade D., Howard Hughes Medical Institute, 660 S. Euclid St. Louis Mo, 63110, USA
REFERENCE 2 (bases 1 to 2376)
AUTHORS Hourcade,D., Miesner,D.R., Atkinson,J.P. and Holers,V.M.
TITLE Identification of an alternative polyadenyiation site in the human C3b/C4b receptor (complement receptor type 1) transcriptional unit and prediction of a secreted form of complement receptor type 1
JOURNAL J. Exp. Med. 168 (4), 1255-1270 (1988)
MEDLINE 89010527
COMMENT The sequence overlaps with that reported by Klickstein et. al. in J. Exp. Med. 165:1095-1112(1987) x05309 and in J. Exp. Med. 168:1699-1717(1988).

FEATURES Location/Qualifiers

source 1. .2376
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CR1-4"
/haplotype="CR1-A"
/cell_line="HL-60"

CDS <1. .1682
/note="CR1 receptor precursor (AA -16 to 543) (1 is 2nd base in codon)"
/codon_start=3
/protein_id="CAA32541.1"
/db_xref="GI:736240"
/translation="SLLAVVVLLALPVAWGQCNAPEWLPFARPTNLTDEFEFPIGTYL
NYECRPGYSGRPFSIICLKNSVWTGAKDRCRRKSCRNPPDPVNGMVHVIKGIQFGSQI
KYSCTKGYRLIGSSSATCIISGDTVIWDNETPICDRIPCGLPPTITNGDFISTNRENF
HYGSVVTYRCNPGSGGRKVFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNV
ENGILVSDNRSLFSLNEVVEFRCQPGFVMKGPRRVKCQALNKWEPELPSCSRVCQPPP
DVLHAERTQRDKDNFSPGQEVFYSCEPGYDLRGAASMRCTPQGDWSPAAPTCEVKSCD
DFMGQLLNGRVLFPVNLQLGAKVDFVCDEGFQLKGSSASYCVLAGMESLWNSSVPVCE
QIFCPSPPVIPNGRHTGKPLEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIRCTSDPQ
GNGVWSSPAPRCGILGHCQAPDHFLFAKLKTQTNASDFPIGTSLKYECRPEYYGRPFS
ITCLDNLVWSSPKDVCKRKSCKTPPDPVNGMVHVITDIQVGSRINYSCTTG"

sig_peptide <1. .50
mat_peptide 51. .1679
/product="mature CR1 receptor (AA 1-543); secreted form"

polyA_site 2376
/note="polyA site"

BASE COUNT 633 a 549 c 568 g 626 t
ORIGIN

alignment_scores:
Quality: 644.50 Length: 136
Ratio: 5.035 Gaps: 1
Percent Similarity: 94.118 Percent Identity: 86.029

alignment_block:
US-10-031-904-8 x HSCR1RS ..

```
Align seg 1/1  to: HSCR1RS  from: 1  to: 2376

     20 LeuLeuAlaAlaLeuValLeuLeu...LeuSerSerPheSerAspGlnCy 35
        |||||||||:::::|||||||||   |||               |||||
      6 CTGCTGGCGGTTGTGGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATG 55

     35 sAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThrAspA 52
        ||||:::|||||||||||||||||||||||||||||||||||||||||||:
     56 CAATGCCCCAGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATG 105

     52 spPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgProGly 68
        ::|||||||||||||||||||||||||||||||||||||||||||||||||
    106 AGTTTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGT 155

     69 TyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerValTrpTh 85
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    156 TATTCCGGAAGACCGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGAC 205

     85 rSerAlaLysAspLysCysLysArgLysSerCysArgAsnProProAspP 102
        |:::|||||||||:::|||:::||||||||||||||||||||||||||||
    206 TGGTGCTAAGGACAGGTGCAGACGTAAATCATGTCGTAATCCTCCAGATC 255

    102 roValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlySerGln 118
        ||||||||||||||:::|||||||||||||   ||||||||||||||||||
    256 CTGTGAATGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAA 305

    119 IleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSerSerAl 135
        ||||||||||||||   ||||||||||||||||||||||||||||||||
    306 ATTAAATATTCTTGTACTAAAGGATACCGACTCATTGGTTCCTCGTCTGC 355

    135 aThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysThrProV 152
        ||||||||||||||||||:::||||||||||||||||||:::|||||||:
    356 CACATGCATCATCTCAGGTGATACTGTCATTTGGGATAATGAAACACCTA 405

    152 alCysAsp 154
        ::|||||||
    406 TTTGTGAC 413

seq_name: gb_sts:G28591

seq_documentation_block:
LOCUS       G28591                 2376 bp    DNA     linear   STS 11-JUL-1996
DEFINITION  human STS SHGC-35372, sequence tagged site.
ACCESSION   G28591
VERSION     G28591.1  GI:1408406
KEYWORDS    STS; STS sequence; primer; sequence tagged site.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2376)
  AUTHORS   Myers,R.M.
  JOURNAL   Unpublished
COMMENT
            Contact: Richard M. Myers
            Stanford Human Genome Center (SHGC)
            Stanford University School of Medicine
            Department of Genetics, M-344, Stanford, CA 94305, USA
            Tel: 4157259687
            Fax: 4157259689
            Email: myers@shgc.stanford.edu

            Primer A: TGAGTTGGCAGCAACATCTC
            Primer B: CATACTCTTATATGTGCACTGCCC
            STS size: 216
            PCR Profile:
                    Initial incubation: 94 degrees C for 90 seconds

                    Denaturation:       94 degrees C for 15 seconds
                    Annealing:          62 degrees C for 23 seconds
                    Polymerization:     72 degrees C for 30 seconds
                    PCR Cycles:         30
                    Thermal Cycler:     Perkin Elmer 9600

            Protocol:
                    Template:           25 ng
                    Primer:             each 1 uM
                    dNTPs:              each 200 uM
                    Taq Polymerase:     0.05 units/ul
                    Total Vol:          10 ul

            Buffer:
                    MgCl2:              2.5 mM
                    KCl:                50  mM
                    Tris-HCl:           20  mM
                    pH:                 8.3

            Prepared with primer pairs provided by Sandoz, derived from X14362
            -- Washington University/Merck EST sequence.
FEATURES             Location/Qualifiers
     source          1. .2376
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="1"
     STS             1680. .1895
     primer_bind     1680. .1699
     primer_bind     complement(1872. .1895)
BASE COUNT      633 a    549 c    568 g    626 t
ORIGIN


alignment_scores:
            Quality:  644.50              Length:    136
              Ratio:  5.035                 Gaps:      1
 Percent Similarity:  94.118   Percent Identity: 86.029

alignment_block:
 US-10-031-904-8 x G28591   ..

 Align seg 1/1  to: G28591  from: 1  to: 2376

      20 LeuLeuAlaAlaLeuValLeuLeu...LeuSerSerPheSerAspGlnCy 35
         |||||||||:::::|||||||||   |||               |||||
       6 CTGCTGGCGGTTGTGGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATG 55

      35 sAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThrAspA 52
         ||||:::|||||||||||||||||||||||||||||||||||||||||||:
      56 CAATGCCCCAGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATG 105

      52 spPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgProGly 68
         ::|||||||||||||||||||||||||||||||||||||||||||||||||
     106 AGTTTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGT 155

      69 TyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerValTrpTh 85
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     156 TATTCCGGAAGACCGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGAC 205

      85 rSerAlaLysAspLysCysLysArgLysSerCysArgAsnProProAspP 102
         |:::|||||||||:::|||:::||||||||||||||||||||||||||||
     206 TGGTGCTAAGGACAGGTGCAGACGTAAATCATGTCGTAATCCTCCAGATC 255

     102 roValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlySerGln 118
         ||||||||||||||:::|||||||||||||   ||||||||||||||||||
     256 CTGTGAATGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAA 305

     119 IleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSerSerAl 135
         ||||||||||||||   ||||||||||||||||||||||||||||||||
     306 ATTAAATATTCTTGTACTAAAGGATACCGACTCATTGGTTCCTCGTCTGC 355

     135 aThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysThrProV 152
         ||||||||||||||||||:::||||||||||||||||||:::|||||||:
     356 CACATGCATCATCTCAGGTGATACTGTCATTTGGGATAATGAAACACCTA 405

     152 alCysAsp 154
         ::|||||||
     406 TTTGTGAC 413
```

```
seq_name: gb_pat:A86593

seq_documentation_block:
LOCUS       A86593                  591 bp    DNA     linear   PAT 21-JAN-2000
DEFINITION  Sequence 2 from Patent WO9839433.
ACCESSION   A86593
VERSION     A86593.1  GI:6735168
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Smith,R.A. and Cox,V.F.
  TITLE     COMPLEMENT RECEPTOR TYPE 1 (CR1)-LIKE SEQUENCES
  JOURNAL   Patent: WO 9839433-A 2 11-SEP-1998;
            SMITH RICHARD ANTONY GODWIN (GB); ADPROTECH PLC (GB)
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      132 a    159 c    148 g    152 t
ORIGIN


  ignment_scores:
            Quality:  635.00             Length:    121
              Ratio:  5.336                Gaps:      0
 Percent Similarity:  98.347   Percent Identity: 90.909

alignment_block:
 US-10-031-904-8 x A86593    ..

 Align seg 1/1  to: A86593  from: 1  to: 591

      34 GlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuTh 50
         |||||||||:::||||||||||||||||||||||||||||||||||||||
       4 CAGTGCAACGCTCCGGAATGGCTGCCGTTCGCGCGCCCGACCAACCTGAC 53

      50 rAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgP 67
         ||||:::||||||||||||||||||||||||||||||||||||||||||||
      54 TGATGAATTTGAGTTCCCGATCGGTACCTACCTGAACTACGAATGCCGCC 103

      67 roGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVal 83
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     104 CGGGTTATAGCGGCCGCCCGTTTTCTATCATCTGCCTGAAAAACTCTGTC 153

      84 TrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProPr 100
         ||||||:::|||||||||:::|||:::||||||||||||||||||||||||
     154 TGGACTGGTGCTAAGGACCGTTGCCGACGTAAATCTTGTCGTAATCCGCC 203

     100 oAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlyS 117
         ||||||||||||||||||:::|||||||||||   |||||||||||||
     204 AGATCCGGTTAACGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGTT 253

     117 erGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSer 133
         ||||||||||||||||||   ||||||||||||||||||||||||||||
     254 CCCAAATTAAATATTCTTGTACTAAAGGTTACCGTCTGATTGGTTCCTCC 303

     134 SerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysTh 150
         ||||||||||||||||||||||:::|||||||||||||||||||:::||
     304 AGCGCTACATGCATCATCTCTGGTGATACTGTCATTTGGGATAATGAAAC 353

     150 rProValCysAsp 154
         ||||:::|||||||
     354 ACCGATTTGTGAC 366

seq_name: gb_pat:A86601

seq_documentation_block:
LOCUS       A86601                  591 bp    DNA     linear   PAT 21-JAN-2000
DEFINITION  Sequence 10 from Patent WO9839433.
ACCESSION   A86601
VERSION     A86601.1  GI:6735175
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Smith,R.A. and Cox,V.F.
  TITLE     COMPLEMENT RECEPTOR TYPE 1 (CR1)-LIKE SEQUENCES
  JOURNAL   Patent: WO 9839433-A 10 11-SEP-1998;
            SMITH RICHARD ANTONY GODWIN (GB); ADPROTECH PLC (GB)
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      127 a    159 c    151 g    154 t
ORIGIN


alignment_scores:
            Quality:  635.00             Length:    121
              Ratio:  5.336                Gaps:      0
 Percent Similarity:  98.347   Percent Identity: 90.909

alignment_block:
 US-10-031-904-8 x A86601    ..

 Align seg 1/1  to: A86601  from: 1  to: 591

      34 GlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuTh 50
         |||||||||:::||||||||||||||||||||||||||||||||||||||
       4 CAGTGCAACGCTCCGGAATGGCTGCCGTTCGCGCGCCCGACCAACCTGAC 53

      50 rAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgP 67
         ||||:::||||||||||||||||||||||||||||||||||||||||||||
      54 TGATGAATTTGAGTTCCCGATCGGTACCTACCTGAACTACGAATGCCGCC 103

      67 roGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVal 83
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     104 CGGGTTATAGCGGCCGCCCGTTTTCTATCATCTGCCTGAAAAACTCTGTC 153

      84 TrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProPr 100
         ||||||:::|||||||||:::|||:::||||||||||||||||||||||||
     154 TGGACTGGTGCTAAGGACCGTTGCCGACGTAAATCTTGTCGTAATCCGCC 203

     100 oAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlyS 117
         ||||||||||||||||||:::|||||||||||   |||||||||||||
     204 AGATCCGGTTAACGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGTT 253

     117 erGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSer 133
         ||||||||||||||||||   ||||||||||||||||||||||||||||
     254 CCCAAATTAAATATTCTTGTACTAAAGGTTACCGTCTGATTGGTTCCTCC 303

     134 SerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysTh 150
         ||||||||||||||||||||||:::|||||||||||||||||||:::||
     304 AGCGCTACATGCATCATCTCTGGTGATACTGTCATTTGGGATAATGAAAC 353

     150 rProValCysAsp 154
         ||||:::|||||||
     354 ACCGATTTGTGAC 366

seq_name: gb_pat:A86603

seq_documentation_block:
LOCUS       A86603                  591 bp    DNA     linear   PAT 21-JAN-2000
DEFINITION  Sequence 12 from Patent WO9839433.
ACCESSION   A86603
VERSION     A86603.1  GI:6735176
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Smith,R.A. and Cox,V.F.
  TITLE     COMPLEMENT RECEPTOR TYPE 1 (CR1)-LIKE SEQUENCES
```

```
  JOURNAL   Patent: WO 9839433-A 12 11-SEP-1998;
            SMITH RICHARD ANTONY GODWIN (GB); ADPROTECH PLC (GB)
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      134 a    158 c    146 g    153 t
ORIGIN


alignment_scores:
           Quality:  635.00             Length:    121
             Ratio:  5.336                Gaps:      0
 Percent Similarity:  98.347   Percent Identity: 90.909

alignment_block:
US-10-031-904-8 x A86603   ..

 Align seg 1/1  to: A86603  from: 1  to: 591

      34 GlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuTh 50
         |||||||||:::||||||||||||||||||||||||||||||||||||||
       4 CAGTGCAACGCTCCGGAATGGCTGCCGTTCGCGCGCCCGACCAACCTGAC 53

      50 rAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgP 67
         ||||:::|||||||||||||||||||||||||||||||||||||||||||
      54 TGATGAATTTGAGTTCCCGATCGGTACCTACCTGAACTACGAATGCCGCC 103

      67 roGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVal 83
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     104 CGGGTTATAGCGGCCGCCCGTTTTCTATCATCTGCCTGAAAAACTCTGTC 153

      84 TrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProPr 100
         ||||||:::||||||||||:::|||:::|||||||||||||||||||||||
     154 TGGACTGGTGCTAAGGACCGTTGCCGACGTAAATCTTGTCGTAATCCGCC 203

     100 oAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlyS 117
         ||||||||||||||||||:::||||||||||||   ||||||||||||||
     204 AGATCCGGTTAACGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGTT 253

     117 erGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSer 133
         ||||||||||||||||||||   ||||||||||||||||||||||||||||
     254 CCCAAATTAAATATTCTTGTACTAAAGGTTACCGTCTGATTGGTTCCTCC 303

     134 SerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysTh 150
         ||||||||||||||||||||||||:::||||||||||||||||||||:::||
     304 AGCGCTACATGCATCATCTCTGGTGATACTGTCATTTGGGATAATGAAAC 353

     150 rProValCysAsp 154
         ||||:::||||||
     354 ACCGATTTGTGAC 366

seq_name: gb_pat:A86605

seq_documentation_block:
LOCUS       A86605                  591 bp    DNA     linear   PAT 21-JAN-2000
DEFINITION  Sequence 14 from Patent WO9839433.
ACCESSION   A86605
VERSION     A86605.1  GI:6735177
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Smith,R.A. and Cox,V.F.
  TITLE     COMPLEMENT RECEPTOR TYPE 1 (CR1)-LIKE SEQUENCES
  JOURNAL   Patent: WO 9839433-A 14 11-SEP-1998;
            SMITH RICHARD ANTONY GODWIN (GB); ADPROTECH PLC (GB)
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      131 a    160 c    149 g    151 t
ORIGIN


alignment_scores:
           Quality:  635.00             Length:    121
             Ratio:  5.336                Gaps:      0
 Percent Similarity:  98.347   Percent Identity: 90.909

alignment_block:
US-10-031-904-8 x A86605   ..

 Align seg 1/1  to: A86605  from: 1  to: 591

      34 GlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuTh 50
         |||||||||:::||||||||||||||||||||||||||||||||||||||
       4 CAGTGCAACGCTCCGGAATGGCTGCCGTTCGCGCGCCCGACCAACCTGAC 53

      50 rAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgP 67
         ||||:::|||||||||||||||||||||||||||||||||||||||||||
      54 TGATGAATTTGAGTTCCCGATCGGTACCTACCTGAACTACGAATGCCGCC 103

      67 roGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVal 83
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     104 CGGGTTATAGCGGCCGCCCGTTTTCTATCATCTGCCTGAAAAACTCTGTC 153

      84 TrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProPr 100
         ||||||:::||||||||||:::|||:::|||||||||||||||||||||||
     154 TGGACTGGTGCTAAGGACCGTTGCCGACGTAAATCTTGTCGTAATCCGCC 203

     100 oAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlyS 117
         ||||||||||||||||||:::||||||||||||   ||||||||||||||
     204 AGATCCGGTTAACGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGTT 253

     117 erGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSer 133
         ||||||||||||||||||||   ||||||||||||||||||||||||||||
     254 CCCAAATTAAATATTCTTGTACTAAAGGTTACCGTCTGATTGGTTCCTCC 303

     134 SerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysTh 150
         ||||||||||||||||||||||||:::||||||||||||||||||||:::||
     304 AGCGCTACATGCATCATCTCTGGTGATACTGTCATTTGGGATAATGAAAC 353

     150 rProValCysAsp 154
         ||||:::||||||
     354 ACCGATTTGTGAC 366

seq_name: gb_pat:A86607

seq_documentation_block:
LOCUS       A86607                  591 bp    DNA     linear   PAT 21-JAN-2000
DEFINITION  Sequence 16 from Patent WO9839433.
ACCESSION   A86607
VERSION     A86607.1  GI:6735178
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Smith,R.A. and Cox,V.F.
  TITLE     COMPLEMENT RECEPTOR TYPE 1 (CR1)-LIKE SEQUENCES
  JOURNAL   Patent: WO 9839433-A 16 11-SEP-1998;
            SMITH RICHARD ANTONY GODWIN (GB); ADPROTECH PLC (GB)
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      128 a    160 c    151 g    152 t
ORIGIN


alignment_scores:
           Quality:  635.00             Length:    121
             Ratio:  5.336                Gaps:      0
 Percent Similarity:  98.347   Percent Identity: 90.909
```

```
alignment_block:
 US-10-031-904-8 x A86607   ..

 Align seg 1/1  to: A86607  from: 1  to: 591

      34 GlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuTh 50
         ||||||||||:::||||||||||||||||||||||||||||||||||||||
       4 CAGTGCAACGCTCCGGAATGGCTGCCGTTCGCGCGCCCGACCAACCTGAC 53

      50 rAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgP 67
         ||||:::||||||||||||||||||||||||||||||||||||||||||||
      54 TGATGAATTTGAGTTCCCGATCGGTACCTACCTGAACTACGAATGCCGCC 103

      67 roGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVal 83
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     104 CGGGTTATAGCGGCCGCCCGTTTTCTATCATCTGCCTGAAAAACTCTGTC 153

      84 TrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProPr 100
         ||||||:::||||||||:::|||:::||||||||||||||||||||||||||
     154 TGGACTGGTGCTAAGGACCGTTGCCGACGTAAATCTTGTCGTAATCCGCC 203

     100 oAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlyS 117
         ||||||||||||||||||:::||||||||||||   ||||||||||||||
     204 AGATCCGGTTAACGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGTT 253

     117 erGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSer 133
         ||||||||||||||||||   ||||||||||||||||||||||||||||||
     254 CCCAAATTAAATATTCTTGTACTAAAGGTTACCGTCTGATTGGTTCCTCC 303

     134 SerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysTh 150
         ||||||||||||||||||||||||:::||||||||||||||||||:::||
     304 AGCGCTACATGCATCATCTCTGGTGATACTGTCATTTGGGATAATGAAAC 353

     150 rProValCysAsp 154
         ||||:::||||||
     354 ACCGATTTGTGAC 366

seq_name: gb_pat:A86609

seq_documentation_block:
LOCUS       A86609                  591 bp    DNA     linear   PAT 21-JAN-2000
DEFINITION  Sequence 18 from Patent WO9839433.
ACCESSION   A86609
VERSION     A86609.1  GI:6735179
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Smith,R.A. and Cox,V.F.
  TITLE     COMPLEMENT RECEPTOR TYPE 1 (CR1)-LIKE SEQUENCES
  JOURNAL   Patent: WO 9839433-A 18 11-SEP-1998;
            SMITH RICHARD ANTONY GODWIN (GB); ADPROTECH PLC (GB)
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      135 a    159 c    146 g    151 t
ORIGIN


alignment_scores:
            Quality:  635.00             Length:    121
              Ratio:  5.336                Gaps:      0
 Percent Similarity:  98.347   Percent Identity: 90.909

alignment_block:
 US-10-031-904-8 x A86609   ..

 Align seg 1/1  to: A86609  from: 1  to: 591

      34 GlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuTh 50
         ||||||||||:::||||||||||||||||||||||||||||||||||||||
       4 CAGTGCAACGCTCCGGAATGGCTGCCGTTCGCGCGCCCGACCAACCTGAC 53

      50 rAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgP 67
         ||||:::||||||||||||||||||||||||||||||||||||||||||||
      54 TGATGAATTTGAGTTCCCGATCGGTACCTACCTGAACTACGAATGCCGCC 103

      67 roGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVal 83
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     104 CGGGTTATAGCGGCCGCCCGTTTTCTATCATCTGCCTGAAAAACTCTGTC 153

      84 TrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProPr 100
         ||||||:::||||||||||:::|||:::||||||||||||||||||||||||||
     154 TGGACTGGTGCTAAGGACCGTTGCCGACGTAAATCTTGTCGTAATCCGCC 203

     100 oAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlyS 117
         ||||||||||||||||||:::||||||||||||   ||||||||||||||
     204 AGATCCGGTTAACGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGTT 253

     117 erGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSer 133
         ||||||||||||||||||   ||||||||||||||||||||||||||||||
     254 CCCAAATTAAATATTCTTGTACTAAAGGTTACCGTCTGATTGGTTCCTCC 303

     134 SerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysTh 150
         ||||||||||||||||||||||||:::||||||||||||||||||:::||
     304 AGCGCTACATGCATCATCTCTGGTGATACTGTCATTTGGGATAATGAAAC 353

     150 rProValCysAsp 154
         ||||:::||||||
     354 ACCGATTTGTGAC 366

seq_name: gb_pat:AR029199

seq_documentation_block:
LOCUS       AR029199                605 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 33 from patent US 5859223.
ACCESSION   AR029199
VERSION     AR029199.1  GI:5941172
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 605)
  AUTHORS   Mossakowska,D.Ewa.Irena., Dodd,I., Freeman,A.Mary. and
            Smith,R.Anthony.Godwin.
  TITLE     Soluble CR1 derivatives
  JOURNAL   Patent: US 5859223-A 33 12-JAN-1999;
FEATURES             Location/Qualifiers
     source          1. .605
                     /organism="unknown"
BASE COUNT      136 a    161 c    150 g    158 t
ORIGIN


alignment_scores:
            Quality:  635.00             Length:    121
              Ratio:  5.336                Gaps:      0
 Percent Similarity:  98.347   Percent Identity: 90.909

alignment_block:
 US-10-031-904-8 x AR029199   ..

 Align seg 1/1  to: AR029199  from: 1  to: 605

      34 GlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuTh 50
         ||||||||||:::||||||||||||||||||||||||||||||||||||||
       7 CAGTGCAACGCTCCGGAATGGCTGCCGTTCGCGCGCCCGACCAACCTGAC 56

      50 rAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgP 67
         ||||:::||||||||||||||||||||||||||||||||||||||||||||
      57 TGATGAATTTGAGTTCCCGATCGGTACCTACCTGAACTACGAATGCCGCC 106

      67 roGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVal 83
```

```
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 107 CGGGTTATAGCGGCCGCCCGTTTTCTATCATCTGCCTGAAAAACTCTGTC 156

  84 TrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProPr 100
     ||||||:::||||||||||:::|||:::|||||||||||||||||||||
 157 TGGACTGGTGCTAAGGACCGTTGCCGACGTAAATCTTGTCGTAATCCGCC 206

 100 oAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlyS 117
     ||||||||||||||||||:::||||||||||||   ||||||||||||
 207 AGATCCGGTTAACGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGTT 256

 117 erGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSer 133
     ||||||||||||||||||   ||||||||||||||||||||||||||||
 257 CCCAAATTAAATATTCTTGTACTAAAGGTTACCGTCTGATTGGTTCCTCC 306

 134 SerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysTh 150
     ||||||||||||||||||||||:::||||||||||||||||||:::||
 307 AGCGCTACATGCATCATCTCTGGTGATACTGTCATTTGGGATAATGAAAC 356

 150 rProValCysAsp 154
     ||||:::||||||
 357 ACCGATTTGTGAC 369

seq_name: gb_pr:BABCORE

seq_documentation_block:
LOCUS       BABCORE                 1688 bp    mRNA    linear   PRI 07-MAY-1996
DEFINITION  Papio cynocephalus complement receptor mRNA, partial cds.
ACCESSION   L77977
VERSION     L77977.1  GI:1301608
KEYWORDS    complement C3b; complement receptor;
            glycophosphatidylinositol-linked protein.
SOURCE      Papio cynocephalus cDNA to mRNA.
  ORGANISM  Papio cynocephalus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Papio.
REFERENCE   1  (bases 1 to 1688)
  AUTHORS   Birmingham,D.J., Logar,C.M., Shen,X.-P. and Chen,W.
  TITLE     The baboon erythrocyte complement receptor is a glycophosphatidyl
            inositol-linked protein encoded by a homologue of the human
            CR1-like genetic element
  JOURNAL   Unpublished (1996)
FEATURES             Location/Qualifiers
     source          1. .1688
                     /organism="Papio cynocephalus"
                     /db_xref="taxon:9556"
                     /cell_type="erythrocyte"
                     /tissue_type="bone marrow"
                     /dev_stage="adult"
     mRNA            <1. .>1688
     CDS             <1. .1571
                     /note="homologue of human CR1-like genetic element"
                     /codon_start=3
                     /product="complement receptor"
                     /protein_id="AAA99004.1"
                     /db_xref="GI:1301609"
                     /translation="VVLQRMGASSPRSPEPVGPPAPRLLFCCGGSLLAVVVLLALPVA
                     WGQCNAPEQLPFARPTNLTDASEFPVGTYLKYECLPGYHGKPFSIICLKNSVWTSAKD
                     KCTRKSCRNPKDPVNGMVHVIKDIQFGSQINYSCNKGYRLIGSSSATCIISGNTVIWD
                     NETPICEIIPCGLPPTIANGDFISTSREYFPYGSVVTYRCNLGSGRKKLFELVGEPSI
                     YCTSKDDQVGIWSGPAPQCIIPNKCMPPNVENGVLVSVNRSLFSLNEVVEFRCQPGFV
                     MKGPRHVQCQALNKWEPELPSCSRVCQPPPEILHGEHTPSHQDFSPGQEVFYSCEPGY
                     DLRGAASLHCTPQGDWNPEAPICTVKSCDDFLGQLPHGRVLFPLNLQLGAKVSFVCDE
                     GFRLKGRFASHCVLAGMKALWNSSVPVCEQIFCPNPPAILNGRHIGAPLGDIPYGKEV
                     SYICDPHPDRGMTVNLIGESTIRCTSDPQGNGVWSSPAPRCELSVPAGANDALIVGTL
                     SGTIFVILFIIFLS"
BASE COUNT      413 a    427 c    411 g    437 t
ORIGIN


alignment_scores:
            Quality:  599.50            Length:    138
```

```
            Ratio:  4.874               Gaps:      1
 Percent Similarity:  89.130   Percent Identity: 81.159

alignment_block:
 US-10-031-904-8 x BABCORE   ..

 Align seg 1/1  to: BABCORE  from: 1  to: 1688

  18 GlyLeuLeuLeuAlaAlaLeuValLeuLeu...LeuSerSerPheSerAs 33
     |||   |||||||||:::::|||||||||   |||
  90 GGATCCCTGTTGGCGGTTGTGGTGCTGCTCGCGCTGCCGGTGGCCTGGGG 139

  33 pGlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuT 50
      ||||||||:::|||||||   ||||||||||||||||||||||||||||
 140 TCAATGCAATGCCCCGGAACAGCTTCCATTTGCCAGGCCTACCAACCTAA 189

  50 hrAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArg 66
     |||||      |||||||||:::|||||||||||||:::|||||||||
 190 CTGATGCATCTGAGTTTCCCGTTGGGACATATCTGAAGTATGAATGCCTC 239

  67 ProGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVa 83
     |||||||||   |||:::||||||||||||||||||||||||||||||||
 240 CCTGGTTATCATGGAAAACCATTTTCTATCATCTGCCTAAAAAACTCAGT 289

  83 lTrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProP 100
     ||||||||||||||||||||||||   ||||||||||||||||||||
 290 CTGGACAAGTGCTAAGGACAAGTGCACACGTAAATCATGTCGTAATCCTA 339

 100 roAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGly 116
       ||||||||||||||||||:::||||||||||||||||||||||||||||
 340 AAGATCCTGTGAATGGCATGGTGCATGTGATCAAAGACATCCAGTTCGGA 389

 117 SerGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSe 133
     |||||||||:::|||||||||   ||||||||||||||||||||||||||
 390 TCCCAAATTAATTATTCTTGTAATAAAGGATACCGACTCATTGGTTCCTC 439

 133 rSerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysT 150
     ||||||||||||||||||||||||||||||||||||||||||||||:::|
 440 GTCCGCCACATGCATCATCTCAGGCAATACTGTCATTTGGGATAATGAAA 489

 150 hrProValCysAsp 154
     |||||:::|||:::
 490 CACCTATTTGTGAG 503

seq_name: gb_pr:BABCR1A

seq_documentation_block:
LOCUS       BABCR1A                 6000 bp    mRNA    linear   PRI 11-FEB-1995
DEFINITION  Papio hamadryas complement component receptor type 1 (CR1) mRNA,
            complete cds.
ACCESSION   L39791
VERSION     L39791.1  GI:662828
KEYWORDS    C3b/C4b complement component receptor; complement component
            receptor CR1; complement receptor 1.
SOURCE      Papio hamadryas (clone SPC-CYT3NC) cDNA to mRNA.
  ORGANISM  Papio hamadryas
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Papio.
REFERENCE   1  (bases 1 to 6000)
  AUTHORS   Clemenza,L., Subramanian,B.V., Nickells,M.W., Hourcade,D.E. and
            Atkinson,J.P.
  TITLE     Primary sequence of the baboon 200 kDa C3b/C4b receptor (CR1)
  JOURNAL   Unpublished (1995)
FEATURES             Location/Qualifiers
     source          1. .6000
                     /organism="Papio hamadryas"
                     /db_xref="taxon:9557"
                     /clone="SPC-CYT3NC"
                     /cell_line="26CB-1"
                     /cell_type="lymphoblastoid"
     gene            1. .6000
                     /gene="CR1"
```

```
    CDS             268. .>6000
                    /gene="CR1"
                    /function="acts as receptor for complement component
                    C3b/C4b"
                    /note="putative"
                    /codon_start=1
                    /product="complement component receptor type 1"
                    /protein_id="AAA62170.1"
                    /db_xref="GI:662829"
                    /translation="MVHVIKDIQFGSQINYSCTEGHRLIGSSSATCIISGNTVIWDNE
                    TPICEKISCGLPPTIDNGDFFSANKEYFHYGSVVTYRCNLGSGGRKLFELVGEPSIYC
                    TSNEDQVGIWSGPAPQCIIPNKCTPPNVENGILVSVNRSLFSLNEVVEFRCQPGFVMK
                    GPRRVQCQALNKWEPELPSCSRVCQPPPDVLHGERTQRDKDIFQTGQEVFYICEPGYD
                    LRGAASLRCTPQGDWSPAAPRCEVKSCDDSLGQLPNGRVLFPRSLQLGAKVDFVCDEG
                    FQLKGSSASYCVLAGMESLWNSSVPVCEQIFCPSPPVIPNGRHTGKPLEVFPFGKAVT
                    YTCDPHPDRGMTFDLIGESTIRCTSDPQGNGVWSSPAPRCGILGHCKAPDHFLFAKLK
                    TQTNASDFPIGTSLKYECRPEYYGRPFSITCLDNLEWSSPKDVCKRKSCKTPPDPVNG
                    MVHVITDIQVGSRINYSCTTGHRLIGHSSAECVTSGNTAHWSTKPPICQRIPCGLPPP
                    IANGDFISTNREYFHYGSVVTYRCNLGSGRKKLFELVGEPSIYCTSKDDQVGIWSGPA
                    PQCIIPNKCMPPNVENGILVSVNRSLFSLNEVVEFRCQPGFVMKGPRRVQCQALNKWE
                    PELPSCSRVCQPPPDVLHGERTQRDKDIFQPGQEVFYICEPGYDLRGAASLRCTPQGD
                    WSPAAPRCEVKSCDDSLGQLPNGRVLFPRSLQLGAKVDFVCDEGFQLKGSSASYCVLA
                    GMESLWNSSVPVCEQIFCPSPPVIPNGRHTGKPLEVFPFGKAVTYTCDPHPDRGMTFD
                    LIGESTIRCTSDPQGNGVWSSRAPRCGILGHCKAPDHFLFAKLKTQTNASDFPIGTSL
                    KYECRPEYYGKPFSITCLDNLVWSSPKDVCKRKSCKTPPDPVNGMVHVITDIQVGSRI
                    NYSCTTGHRLIGHSSAECIISGNTAHWSTKPPICQRIPCGLPPPIANGDFISTNREYF
                    HYGSVVTYRCNLGSGRKKLFELVGEPSIYCTSKDDQVGIWSGPAPQCIIPNKCMPPNV
                    ENGILVSVNRSLFSLNEVVEFRCQPGFVMKGPRRVQCQALNKWEPELPSCSRVCQPPP
                    EILHGEHTPSHQDKFSPGQEVFYSCEPGYDLRGAASLHCTPQGDWSPEAPRCAVKSCD
                    DFLGQLHHGRVLVPFNLQLGAKVSFVCDEGFRLKGSSVSHCVLVGMRSLWNNSVPVCE
                    HIFCPNPPAILNGRHTGALLGDIPYGKEISYTCDPHRDRGMTFNLIGESTIRCTSDLQ
                    GNGVWSSPAPRCELSVRAGHCKTPEQFPFASPTIPINDFEFPVGTSLNYECHPGYFGR
                    MFSISCLENLVWSSVEDNCRRKSCGTPPEPFNGMVHINTDTQFGSTVNYSCNEGFRLI
                    GSPSTTCLVSGNNVTWDKEAPICEIISCKPPPTISNGDFYSNNRTSFHSGTVVTYQCH
                    TGPDGEQLFELVGERSIYCTSKDDQVGAWSSPPPRCISTNKCTAPEVKNGIRVPGNRS
                    FFSLNEIVRFRCQPGFVMVGSHTVQCQTNNRWGPKLPHCSRVCQPPPEILHGEHTPSH
                    QDKFSPGQEVFYSCEPGYDLRGAASLHCTPQGDWSPEAPICTVKSCDEFLGQLPHGRV
                    LSPLNLQLGAKVSFVCDEGFRLKGRSASHCVLAGMKALWNSSVPVCEQIFCPNPPAIL
                    NGRHTGTPLGDIPYGKEVSYTCDPHPDRGMTFNLIGESTIRCTSDLQGNGVWSSPAPR
                    CELSVPAACPHPPKIQNGHYIGGHVSLYLPGMTIGYICDPGYLLVGKGIIFCTDQGIW
                    SQLDHYCKEVNCSFPQFMNGISKELEMKKVYHYGDYVTLECEDGYALEGSPWSQCQAD
                    DRWDPPLAICTSRARDALIVGTLSGMIFVILFIIFLSWIILKYRKGNNAHEKPKEVAI
                    HIHSQGGSSVH"
BASE COUNT     1540 a   1461 c   1465 g   1534 t
ORIGIN


alignment_scores:
            Quality:  575.50             Length:    138
              Ratio:  4.679                Gaps:      1
 Percent Similarity:  89.130   Percent Identity: 78.986

Alignment_block:
US-10-031-904-8 x BABCR1A   ..

 Align seg 1/1  to: BABCR1A  from: 1  to: 6000

     18 GlyLeuLeuLeuAlaAlaLeuValLeuLeu...LeuSerSerPheSerAs 33
        |||   |||||||||:::::|||||||||   |||
      1 GGATCCCTGCTGGCGGTTGTGGTGCTGCTCGCGCTGCCGGTGGCCTGGGG 50

     33 pGlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuT 50
         ||||||||||:::||||||   |||||||||||||||||||:::|||
     51 TCAATGCAATGCCCCGGAACAGCTTCCATTTGCCAGGCCTACTGAACTAA 100

     50 hrAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArg 66
          |||:::   ||||||   ||||||||||:::|||:::||||||||||||
    101 TTGATGAGCCTGAGTTTTCCATTGGGACACATCTGAAGTATGAATGTCGC 150

     67 ProGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVa 83
        |||||||||   ||||||||||||||||||||||||||||||||||||||
    151 CCTGGTTATTATGGAAGACCATTTTCTATCATCTGCCTAAAAAACTCAGT 200

     83 lTrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProP 100
        ||||||||||||||||||||||||   ||||||||||||||||||||||
    201 CTGGACAAGTGCTAAGGACAAGTGCATACGTAAATCATGTCGTAATCCTA 250

    100 roAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGly 116
          ||||||||||||||||||:::|||||||||||||||||||||||||||
    251 GAGATCCTGTGAATGGCATGGTGCATGTGATCAAAGACATCCAGTTCGGA 300

    117 SerGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSe 133
        |||||||||:::|||||||||   :::|||:::||||||||||||||||||
    301 TCCCAGATTAATTATTCTTGTACTGAAGGACACCGACTCATTGGTTCCTC 350

    133 rSerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysT 150
        ||||||||||||||||||||||||||||||||||||||||||||||:::|
    351 ATCTGCCACATGCATCATCTCAGGCAATACTGTCATTTGGGATAATGAAA 400

    150 hrProValCysAsp 154
        |||||:::|||:::
    401 CACCTATTTGTGAG 414
```

OM of: US-10-031-904-8 to: EST:* out_format : pfs

Date: Oct 9, 2002 6:07 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10031904/runat_09102002_084256_23627/app_query.fasta_1.232
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=2000000000 -USER=US10031904_@CGN1_1_3787
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
[redacted]NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-10-031-904-8
Query length: 174
Database: EST:*
Database sequences: 13736207
Database length: -1841457050
Search time (sec): 1598.320000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_est2:BE552138	-	789.00	1553.80	2.7e-77	465	BE552138 hw29d02.x1 NCI_CGAP_Ki
gb_est1:AI240881	-	748.00	1472.86	8.6e-73	443	AI240881 qj94e04.x1 NCI_CGAP_Ki
gb_est1:AI718588	-	699.00	1375.52	2.3e-67	444	AI718588 as46h01.x1 Barstead ac
gb_est2:BF240184	+	678.00	1327.40	1.1e-64	832	BF240184 601905704F1 NIH_MGC_54
gb_est2:BM477528	+	649.50	1268.30	2.1e-61	1063	BM477528 AGENCOURT_6484929 NIH
gb_est1:AI735459	-	603.00	1187.10	7.1e-57	356	AI735459 at11h03.x1 Barstead ac
gb_est2:T66824	+	555.00	1092.05	1.4e-51	346	T66824 ya50b08.r3 Soares fetal
gb_est2:T66823	-	551.50	1083.91	4.0e-51	389	T66823 ya50b08.s3 Soares fetal
gb_est2:T83269	+	532.00	1041.82	8.8e-49	541	T83269 yd41a11.r1 Soares fetal
gb_est2:H73873	+	497.00	974.41	5.0e-45	440	H73873 ys14d08.r1 Soares fetal
gb_est1:AA212152	+	420.50	819.29	2.2e-36	601	AA212152 mu80g07.r1 Stratagene
gb_est2:BG077250	+	418.50	815.73	3.4e-36	577	BG077250 H3013A12-5 NIA Mouse 1
gb_est2:BI697900	+	418.50	814.39	4.1e-36	658	BI697900 603344033F1 NCI_CGAP_M
gb_est2:BI455761	+	418.50	812.47	5.2e-36	795	BI455761 603171769F1 NCI_CGAP_M
[redacted]_est2:BI078162	+	418.50	812.31	5.3e-36	807	BI078162 602872480F1 NCI_CGAP_M
[redacted]_est1:AI735085	-	412.00	805.75	1.2e-35	433	AI735085 as44f09.x1 Barstead ac
gb_est1:AA107525	+	410.50	799.21	2.9e-35	614	AA107525 mp08b07.r1 Life Tech m
gb_est1:BB613937	+	407.50	792.15	7.1e-35	684	BB613937 RIKEN full-le
gb_htc:AK004825	+	405.50	779.23	3.7e-34	1646	AK004825 Mus musculus adult ma
gb_est2:BF023537	+	405.00	790.63	8.6e-35	488	BF023537 ux09c08.y1 Soares_thym
gb_est2:BI693880	+	405.00	789.06	1.1e-34	569	BI693880 603342124F1 NCI_CGAP_M
gb_est2:BI696340	+	404.50	785.64	1.6e-34	722	BI696340 603345540F1 NCI_CGAP_M
gb_est2:BI556660	+	404.00	783.95	2.0e-34	773	BI556660 603239267F1 NIH_CGAP_M
gb_est1:AA286570	+	401.00	781.94	2.6e-34	525	AA286570 vb76a03.r1 Soares mous
gb_est1:AW610808	+	392.00	761.89	3.4e-33	650	AW610808 up37f12.x1 Soares_mous
gb_est1:AW610947	+	385.50	753.85	9.6e-33	403	AW610947 un74b11.y1 Sugano mous
gb_est1:AW012541	+	385.50	753.70	9.8e-33	409	AW012541 uo06c01.y1 Sugano mous
gb_est1:AW912091	+	385.50	749.09	1.8e-32	643	AW912091 uf41f04.y1 Soares_mamm
gb_est1:AW825298	+	384.00	749.95	1.6e-32	441	AW825298 us13d06.y1 Soares_NMGE
gb_est2:T27695	+	382.50	749.18	1.8e-32	355	T27695 EST12512 Human Uterus Hc
gb_est2:BE917021	+	378.00	729.88	2.1e-31	981	BE917021 601663878F1 NCI_CGAP_M
gb_est2:BG146321	+	377.00	737.35	8.0e-32	388	BG146321 mab90c07.y1 NCI_CGAP_S
gb_est2:BF178335	+	376.00	727.24	2.9e-31	861	BF178335 601808574F1 NCI_CGAP_M
gb_est1:AA261275	+	371.50	724.50	4.2e-31	469	AA261275 vb03b10.r1 Soares mous
gb_est2:BG243080	+	370.50	717.88	9.7e-31	739	BG243080 602355374F1 NCI_CGAP_M
gb_est1:AW211552	+	356.00	695.67	1.7e-29	387	AW211552 uo82e02.y1 NCI_CGAP_Ma
gb_est2:BF166939	+	350.00	675.28	2.3e-28	889	BF166939 601774087F1 NCI_CGAP_L
gb_est2:BF015454	+	340.50	663.79	1.0e-27	431	BF015454 uy21g02.y1 NCI_CGAP_Lu
gb_est2:BF159845	+	330.00	641.04	1.8e-26	519	BF159845 601767506F1 NCI_CGAP_L
gb_est2:BF124847	+	328.50	634.04	4.5e-26	770	BF124847 601756853F1 NCI_CGAP_M
gb_htc:AK017702	+	323.50	607.68	1.3e-24	3859	AK017702 Mus musculus 8 day
gb_est1:AW919879	+	314.00	608.69	1.2e-24	549	AW919879 EST351183 Rat gene
gb_est1:AA756782	+	309.50	602.51	2.6e-24	419	AA756782 vv65f04.r1 Stratage
gb_est2:BM488775	+	289.50	559.65	6.3e-22	570	BM488775 pgm2n.pk008.m6 Norm
gb_est2:BM425702	+	289.50	558.90	6.9e-22	613	BM425702 pgf2c.pk001.j17 Pri

seq_name: gb_est2:BE552138

seq_documentation_block:
LOCUS BE552138 465 bp mRNA linear EST 10-AUG-2000
DEFINITION hw29d02.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3184323 3' similar to gb:Y00816_cds1 COMPLEMENT RECEPTOR TYPE 1 PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION BE552138
VERSION BE552138.1 GI:9793830
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 465)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 456.
FEATURES Location/Qualifiers
source 1. .465
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3184323"
/clone_lib="NCI_CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 134 a 85 c 107 g 139 t
ORIGIN

alignment_scores:
Quality: 789.00 Length: 142
Ratio: 5.556 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-10-031-904-8 x BE552138/rev ..

Align seg 1/1 to reverse of: BE552138 from: 1 to: 465

```
     33 AspGlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLe 49
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    453 GATCAATGCAATGTCCCGGAATGGCTTCCATTTGCCAGGCCTACCAACCT 404

     49 uThrAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysA 66
```

```
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
    403 AACTGATGACTTTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCC 354

     66 rgProGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSer 82
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    353 GCCCTGGTTATTCCGGAAGACCGTTTTCTATCATCTGCCTAAAAAACTCA 304

     83 ValTrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnPr 99
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    303 GTCTGGACAAGTGCTAAGGACAAGTGCAAACGTAAATCATGTCGTAATCC 254

     99 oProAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheG 116
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    253 TCCAGATCCTGTGAATGGCATGGCACATGTGATCAAAGACATCCAGTTCG 204

    116 lySerGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySer 132
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    203 GATCCCAAATTAAATATTCTTGTCCTAAAGGATACCGACTCATTGGTTCC 154

    133 SerSerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLy 149
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    153 TCGTCTGCCACATGCATCATCTCAGGCAACACTGTCATTTGGGATAATAA 104

    149 sThrProValCysAspSerGluLeuLysTyrAlaPheLeuPheLeuLeuP 166
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    103 AACACCTGTTTGTGACAGTGAGTTGAAATATGCATTCCTATTTCTTTTAC 54

    166 roIleHisSerAsnPheSerLeuGlu 174
        |||||||||||||||||||||||||
     53 CGATACATTCTAATTTTTCTCTGGAA 28

seq_name: gb_est1:AI240881

seq_documentation_block:
LOCUS       AI240881                443 bp    mRNA    linear   EST 28-JAN-1999
DEFINITION  qj94e04.x1 NCI_CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1867134 3'
            similar to gb:Y00816_cds1 COMPLEMENT RECEPTOR TYPE 1 PRECURSOR
            (HUMAN);, mRNA sequence.
ACCESSION   AI240881
VERSION     AI240881.1  GI:3836278
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 625   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 373.
FEATURES             Location/Qualifiers
     source          1. .443
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1867134"
                     /clone_lib="NCI_CGAP_Kid3"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer,
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not
                     I and Eco RI sites of the modified pT7T3 vector. mRNA
                     source: 2 pooled kidneys. Library went through one round
                     of normalization. Library constructed by Bento Soares and
                     M. Fatima Bonaldo. "
BASE COUNT      131 a     80 c    104 g    128 t
ORIGIN


alignment_scores:
            Quality:  748.00              Length:    141
              Ratio:  5.420                 Gaps:      0
 Percent Similarity:  97.872   Percent Identity: 97.163

alignment_block:
 US-10-031-904-8 x AI240881/rev   ..

 Align seq 1/1  to reverse of: AI240881  from: 1  to: 443

      34 GlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuTh 50
         ||||||||||||   :::   ||||||||||||||||||||||||||||||
    442°CAATGCAATGTCCGGAATGGGCTTCCATTTGCCAGGCCTACCAACCTAAC 393

      50 rAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgP 67
         |||||||   ||||||||||||||||||||||||||||||||||||||||
    392 TGATGACTCTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCC 343

      67 roGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVal 83
         |||||||||||||||||||||||||||||||||||||||||||||||||
    342 CTGGTTATTCCGGAAGACCGTTTTCTATCATCTGCCTAAAAAACTCAGTC 293

      84 TrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProPr 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    292 TGGACAAGTGCTAAGGACAAGTGCAAACGTAAATCATGTCGTAATCCTCC 243

     100 oAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlyS 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    242 AGATCCTGTGAATGGCATGGCACATGTGATCAAAGACATCCAGTTCGGAT 193

     117 erGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSer 133
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    192 CCCAAATTAAATATTCTTGTCCTAAAGGATACCGACTCATTGGTTCCTCG 143

     134 SerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysTh 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    142 TCTGCCACATGCATCATCTCAGGCAACACTGTCATTTGGGATAATAAAAC 93

     150 rProValCysAspSerGluLeuLysTyrAlaPheLeuPheLeuLeuProI 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      92 ACCTGTTTGTGACAGTGAGTTGAAATATGCATTCCTATTTCTTTTACCGA 43

     167 leHisSerAsnPheSerLeuGlu 174
         |||||||||||||||||||||||
      42 TACATTCTAATTTTTCTCTGGAA 20

seq_name: gb_est1:AI718588

seq_documentation_block:
LOCUS       AI718588                444 bp    mRNA    linear   EST 10-JUN-1999
DEFINITION  as46h01.x1 Barstead aorta HPLRB6 Homo sapiens cDNA clone
            IMAGE:2320273 3' similar to gb:Y00816_cds1 COMPLEMENT RECEPTOR TYPE
            1 PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   AI718588
VERSION     AI718588.1  GI:5035844
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 444)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
```

```
          ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
          White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
          Washington University School of Medicine
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
          Tel: 314 286 1800
          Fax: 314 286 1810
          Email: est@watson.wustl.edu
          This clone is available royalty-free through LLNL ; contact the
          IMAGE Consortium (info@image.llnl.gov) for further information.
          Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .444
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2320273"
                     /clone_lib="Barstead aorta HPLRB6"
                     /sex="male"
                     /dev_stage="adult, age 64"
                     /lab_host="DH10B (phage resistant)"
                     /note="Organ: aorta; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [5' AATTCGGATCGAAC 3' and 5' GTTGGATCGG 3'], digested
                     with Not I and cloned into the Not I and Eco RI sites of
                     the modified pT7T3 vector. Library constructed by Bob
                     Barstead."
BASE COUNT      134 a      87 c     103 g     120 t
ORIGIN


alignment_scores:
            Quality:  699.00            Length:    141
              Ratio:  5.102               Gaps:      0
 Percent Similarity:  97.163   Percent Identity: 90.780

alignment_block:
 US-10-031-904-8 x AI718588/rev   ..

 Align seg 1/1  to reverse of: AI718588  from: 1  to: 444

      34 GlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuTh 50
         |||||||||:::|||||||||||||||||||||||||||||||||||||||
     441 CAATGCAATGCCCCAGAATGGCTTCCATTTGCCAGGCCTACCAACCTAAC 392

      50 rAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgP 67
         ||||:::|||||||||||||||||||||||||||||||||||||||||||
     391 TGATGAGTTTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCC 342

      67 roGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVal 83
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     341 CTGGTTATTCCGGAAGACCGTTTTCTATCATCTGCCTAAAAAACTCAGTC 292

      84 TrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProPr 100
         ||||||:::|||||||||:::|||:::||||||||||||||||||||||
     291 TGGACTGGTGCTAAGGACAGGTGCAGACGTAAATCATGTCGTAATCCTCC 242

     100 oAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlyS 117
         ||||||||||||||||||||:::|||||||||||   ||||||||||||||
     241 AGATCCTGTGAATGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGAT 192

     117 erGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSer 133
         ||||||||||||||||||||   ||||||||||||||||||||||||||||
     191 CCCAAATTAAATATTCTTGTACTAAAGGATACCGACTCATTGGTTCCTCG 142

     134 SerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysTh 150
         ||||||||||||||||||||||:::|||||||||||||||||||:::|||
     141 TC.GCCACATGCATCATCTCAGGTGATACTGTCATTTGGGATAATGAAAC 93

     150 rProValCysAspSerGluLeuLysTyrAlaPheLeuPheLeuLeuProI 167
         ||||:::|||||||||||||||||||||||   ||||||||||||||||||
      92 ACCTATTTGTGACAGTGAGTTGAAATATCCCTTCCTATTTCTTTTACCGA 43

     167 leHisSerAsnPheSerLeuGlu 174
         ||||||||||||||||||||||
      42 CCCATTCTAATTTTTCTCTGGAA 20

seq_name: gb_est2:BF240184

seq_documentation_block:
LOCUS       BF240184                832 bp    mRNA    linear   EST 14-NOV-2000
DEFINITION  601905704F1 NIH_MGC_54 Homo sapiens cDNA clone IMAGE:4133424 5',
            mRNA sequence.
ACCESSION   BF240184
VERSION     BF240184.1  GI:11154107
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 832)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1034  row: h  column: 01
            High quality sequence stop: 544.
FEATURES             Location/Qualifiers
     source          1. .832
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4133424"
                     /clone_lib="NIH_MGC_54"
                     /tissue_type="from chronic myelogenous leukemia"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc
                     ); Double-stranded cDNA was prepared from cell line RNA.
                     5' and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).   Average
                     insert size 1.75 kb (range 0.9-4.0 kb).  15/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
BASE COUNT      188 a     226 c     197 g     220 t       1 others
ORIGIN


alignment_scores:
            Quality:  678.00            Length:    163
              Ratio:  4.612               Gaps:      3
 Percent Similarity:  90.184   Percent Identity: 88.957

alignment_block:
 US-10-031-904-8 x BF240184   ..

 Align seg 1/1  to: BF240184  from: 1  to: 832

       1 MetAlaProProValArgLeuGluArgProPheProSerArgArgPhePr 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     113 ATGGCGCCTCCCGTCCGTCTCGAGCGTCCCTTTCCTTCCCGGCGCTTTCC 162
```

```
     17 oGlyLeuLeuLeuAlaAlaLeuValLeuLeuLeuSerSerPheSerAspG 34
        ||||||||||||||||   ||||||||||||||||||||||||||||||||
    163 TGGGTTGCTTCTGGCGGC.CTGGTGTTGCTGCTGTCCTCCTTCTCCGATC 211

     34 lnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThr 50
        |||||||||||||||||||||||||||||||||      |||||||||
    212 AATGCAATGTCCCGGAATGGCTTCCATTTGCCAGGCTACCAAACCTAACT 261

     51 AspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgPr 67
        ||||||||||||||||||||   ||||||||||||||||||||||||||
    262 GATGACTTTGAGTTTCCCATTNGGACATATCTGAACTATGAATGCCGCCC 311

     67 oGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerValT 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    312 TGGTTATTCCGGAAGACCGTTTTCTATCATCTGCCTAAAAAACTCAGTCT 361

     84 rpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProPro 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    362 GGACAAGTGCTAAGGACAAGTGCAAACGTAAATCATGTCGTAATCCTCCA 411

    101 AspProValAsnGlyMetAla.HisValIleLysAspIleGlnPhe.Gly 116
        |||||||||||||||||||| ||||||||||||||||||||||||||
    412 GATCCTGTGAATGGCATGGCACCATGTGATCAAAGACATCCAGTTCAAGA 461

    117 SerGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSe 133
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    462 TCCCAAATTAAATATTCTTGTCCTAAAGGATACCGACTCATTGGTTCCTC 511

    133 rSerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysT 150
        ||||   ||||||||||||||||||||||||||||||
    512 GTCTGC.ACATGCATCATCTCAGGCAACACTGTCATTCTGGGCATACATA 560

    150 hrProValCys......AspSerGluLeuLysTyr 159
           :::|||      ||||||:::      |||
    561 AAACCACCTGTTTTGTGGACAGTGATGACCATTAC 595
```

seq_name: gb_est2:BM477528

seq_documentation_block:

```
LOCUS       BM477528                1063 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6484929 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5554371
            5', mRNA sequence.
ACCESSION   BM477528
VERSION     BM477528.1  GI:18526570
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1063)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12272  row: n  column: 04
            High quality sequence stop: 749.
FEATURES             Location/Qualifiers
     source          1. .1063
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5554371"
                     /clone_lib="NIH_MGC_85"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      251 a    286 c    272 g    249 t      5 others
ORIGIN
```

alignment_scores:

```
            Quality:  649.50             Length:    148
              Ratio:  4.883                Gaps:      2
 Percent Similarity:  89.865   Percent Identity: 81.757
```

alignment_block:

US-10-031-904-8 x BM477528 ..

Align seg 1/1 to: BM477528 from: 1 to: 1063

```
     10 ProPheProSerArgArgPhe......ProGlyLeuLeuLeuAlaAlaLe 24
        |||   |||:::     |||        |||   |||||||||:::::
     61 CCGGCGCCCGGTCTCCCCTTCTGCTGCGGAGGATCCCTGCTGGCGGTTGT 110

     24 uValLeuLeu...LeuSerSerPheSerAspGlnCysAsnValProGluT 40
        :|||||||||   |||              |||||||||:::|||||||
    111 GGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAGAAT 160

     40 rpLeuProPheAlaArgProThrAsnLeuThrAspAspPheGluPhePro 56
        ||||||||||||||||||||||||||||||:::|||||||||||||||
    161 GGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCCC 210

     57 IleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgPr 73
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    211 ATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGACC 260

     73 oPheSerIleIleCysLeuLysAsnSerValTrpThrSerAlaLysAspL 90
        ||||||||||||||||||||||||||||||||||:::|||||||||:
    261 GTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACTGGTGCTAAGGACA 310

     90 ysCysLysArgLysSerCysArgAsnProProAspProValAsnGlyMet 106
        ::|||:::||||||||||||||||||||||||||||||||||||||||||
    311 GGTGCAGACGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATG 360

    107 AlaHisValIleLysAspIleGlnPheGlySerGlnIleLysTyrSerCy 123
        :::|||||||||||||   ||||||||||||||||||||||||||||||
    361 GTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAAATTAAATATTCTTG 410

    123 sProLysGlyTyrArgLeuIleGlySerSerSerAlaThrCysIleIleS 140
        |   ||||||||||||||||||||||||||||||||||||||||||||||
    411 TACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCATCT 460

    140 erGlyAsnThrValIleTrpAspAsnLysThrProValCysAsp 154
        |||||:::|||||||||||||||||:::|||||:::||||||
    461 CAGGTGATACTGTCATTTGGGATAATGAAACACCTATTTGTGAC 504
```

seq_name: gb_est1:AI735459

seq_documentation_block:

```
LOCUS       AI735459                356 bp    mRNA    linear   EST 14-JUN-1999
DEFINITION  at11h03.x1 Barstead aorta HPLRB6 Homo sapiens cDNA clone
            IMAGE:2354837 3' similar to gb:Y00816_cds1 COMPLEMENT RECEPTOR TYPE
            1 PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   AI735459
VERSION     AI735459.1  GI:5056983
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 356)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
```

```
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .356
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2354837"
                     /clone_lib="Barstead aorta HPLRB6"
                     /sex="male"
                     /dev_stage="adult, age 64"
                     /lab_host="DH10B (phage resistant)"
                     /note="Organ: aorta; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [5' AATTCGGATCGAAC 3' and 5' GTTGGATCGG 3'], digested
                     with Not I and cloned into the Not I and Eco RI sites of
                     the modified pT7T3 vector. Library constructed by Bob
                     Barstead."
BASE COUNT      107 a     61 c     77 g    111 t
ORIGIN


alignment_scores:
            Quality:  603.00             Length:    114
              Ratio:  5.432                Gaps:      0
 Percent Similarity:  97.368   Percent Identity: 97.368

alignment_block:
 US-10-031-904-8 x AI735459/rev   ..

 Align seg 1/1  to reverse of: AI735459  from: 1  to: 356

      61 LeuAsnTyrGluCysArgProGlyTyrSerGlyArgProPheSerIleIl 77
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     355 CTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGACCGTTTTCTATCAT 306

      77 eCysLeuLysAsnSerValTrpThrSerAlaLysAspLysCysLysArgL 94
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     305 CTGCCTAAAAAACTCAGTCTGGACAAGTGCTAAGGACAAGTGCAAACGTA 256

      94 ysSerCysArgAsnProProAspProValAsnGlyMetAlaHisValIle 110
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     255 AATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATGGCACATGTGATC 206

     111 LysAspIleGlnPheGlySerGlnIleLysTyrSerCysProLysGlyTy 127
         ||||||||||||||||   ||||||||||||||||||||||||||||||||
     205 AAAGACATCCAGTTCAGATCCCAAATTAAATATTCTTGTCCTAAAGGATA 156

     127 rArgLeuIleGlySerSerSerAlaThrCysIleIleSerGlyAsnThrV 144
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     155 CCGACTCATTGGTTCCTCGTCTGCCACATGCATCATCTCAGGCAACACTG 106

     144 alIleTrpAspAsnLysThrProValCysAspSerGluLeuLysTyrAla 160
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     105 TCATTTGGGATAATAAAACACCTGTTTGTGACAGTGAGTTGAAATATGCA 56

     161 PheLeuPheLeuLeuProIleHisSerAsnPheSerLeuGlu 174
         ||||||||||||       ||||||||||||||||||||||||
      55 TTCCTATTTCTTAAAAAGATACATTCTAATTTTTCTCTGGAA 14

seq_name: gb_est2:T66824

seq_documentation_block:
LOCUS       T66824                 346 bp    mRNA    linear   EST 07-MAR-1995
DEFINITION  ya50b08.r3 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:66327 5' similar to gb:Y00816_cds1 COMPLEMENT RECEPTOR TYPE 1
            PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   T66824
VERSION     T66824.1  GI:676264
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 346)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Other_ESTs: ya50b08.s3.exp
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High qality sequence stops: 324
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: T7
            High quality sequence stop: 324.
FEATURES             Location/Qualifiers
     source          1. .346
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:66327"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       67 a    101 c     77 g     97 t      4 others
ORIGIN


alignment_scores:
            Quality:  555.00             Length:    108
              Ratio:  5.236                Gaps:      2
 Percent Similarity:  98.148   Percent Identity: 97.222

alignment_block:
 US-10-031-904-8 x T66824   ..

 Align seg 1/1  to: T66824  from: 1  to: 346

       4 ProValArgLeuGluArgProPheProSerArgArgPheProGlyLeuLe 20
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      24 CCCGTCCGTCTCGAGCGTCCCTTTCCTTCCCGGCGCTTTCCTGGGTTGCT 73

      20 uLeuAlaAlaLeuValLeuLeuLeuSerSerPheSerAspGlnCysAsnV 37
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      74 TCTGGCGGCCCTGGTGTTGCTGCTGTCCTCCTTCTCCGATCAATGCAATG 123

      37 alProGluTrpLeuProPheAlaArgProThrAsnLeuThrAspAspPhe 53
```

```
      ||||||||||||||||||||||||||||||||||||||||||||||||||
    124 TCCCGGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGACTTT 173

     54 GluPheProIleGlyThrTyrLeuAsnTyrGluCysArgPro.GlyTyrS 70
        |||||||||||||||||||||||||||||||||||||||||||| |||||||
    174 GAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGGTTATT 223

     70 erGlyArgProPheSerIleIleCysLeuLysAsnSerValTrpThrSer 86
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    224 CCGGAAGACCGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACAAGT 273

     87 AlaLysAspLysCysLysArg.LysSerCysArgAsnProProAspProV 103
        |||||||||||||||||||| ||||||||||||||||||||||||||||
    274 GCTAAGGACAAGTGCAAACGTAAAATCATGTCGTAATCCTCCAGATCCTG 323

    103 alAsnGlyMetAlaHisVal 109
        ||||||||||:::||||||
    324 TGAATGGNATGGNACATGTG 343

seq_name: gb_est2:T66823

seq_documentation_block:
LOCUS       T66823                  389 bp    mRNA    linear   EST 07-MAR-1995
DEFINITION  ya50b08.s3 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:66327 3' similar to gb:Y00816_cds1 COMPLEMENT RECEPTOR TYPE 1
            PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   T66823
VERSION     T66823.1  GI:676263
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 389)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Other_ESTs: ya50b08.r3.exp
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -21M13LR
            High quality sequence stop: 390.
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:66327"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      114 a     67 c     92 g    112 t      4 others
ORIGIN
```

```
alignment_scores:
            Quality:  551.50              Length:    123
              Ratio:  4.924                 Gaps:      4
 Percent Similarity:  91.057   Percent Identity: 90.244

alignment_block:
 US-10-031-904-8 x T66823/rev   ..

 Align seg 1/1  to reverse of: T66823  from: 1  to: 389

      56 ProIleGlyThrTyr.LeuAsnTyrGluCysArgPro...GlyTyrSerG 71
         |||:::|||||||| ||||||||||||||||   |||          |
     385 CCATTGGGGACATATCCTGAACTATGAATGCCCGCCCCTGGTTANTCCCG 336

      71 lyArgProPheSer.IleIleCys.LeuLysAsnSerValTrpThrSerA 87
         |||||||||||||| ||||||||| |||||||||||||||||||||||||
     335 GAAGACCGTTTTCCTATCATCTGCCCTAAAAAACTCAGTCTGGACAAGTG 286

      87 laLysAspLysCysLysArgLysSerCysArgAsnProProAspProVal 103
         |||||||||||||||||||||||||||||||   ||||||   ||||||
     285 CTAAGGACAAGTGCAAACGTAAATCATGTCGTAANCCTCCAGANCCTGTG 236

     104 AsnGlyMetAlaHisValIleLysAspIleGlnPheGlySerGlnIleLy 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     235 AATGGCATGGCACATGTGATCAAAGACATCCAGTTCGGATCCCAAATTAA 186

     120 sTyrSerCysProLysGlyTyrArgLeuIleGlySerSerSerAlaThrC 137
         |   |||||||||||||||||||||||||||||||||||||||||||||||
     185 ATANTCTTGTCCTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACAT 136

     137 ysIleIleSerGlyAsnThrValIleTrpAspAsnLysThrProValCys 153
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     135 GCATCATCTCAGGCAACACTGTCATTTGGGATAATAAAACACCTGTTTGT 86

     154 AspSerGluLeuLysTyrAlaPheLeuPheLeuLeuProIleHisSerAs 170
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      85 GACAGTGAGTTGAAATATGCATTCCTATTTCTTTTACCGATACATTCTAA 36

     170 nPheSerLeuGlu 174
         ||||||||||||
      35 TTTTTCTCTGGAA 23

seq_name: gb_est2:T83269

seq_documentation_block:
LOCUS       T83269                  541 bp    mRNA    linear   EST 16-MAR-1995
DEFINITION  yd41a11.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:110780 5' similar to gb:Y00816_cds1 COMPLEMENT RECEPTOR TYPE
            1 PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   T83269
VERSION     T83269.1  GI:711557
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 541)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 791
            High quality sequence stops: 300 Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
```

```
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 791   Std Error: 0.00
            Seq primer: M13RPl
            High quality sequence stop: 300.
FEATURES              Location/Qualifiers
     source           1. .541
                      /organism="Homo sapiens"
                      /db_xref="GDB:466397"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:110780"
                      /clone_lib="Soares fetal liver spleen 1NFLS"
                      /sex="male"
                      /dev_stage="20 week-post conception fetus"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                      with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                      1st strand cDNA was primed with a Pac I - oligo(dT) primer
                      [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Pac I and cloned into the Pac I
                      and Eco RI sites of the modified pT7T3 vector.  Library
                      went through one round of normalization. Library
                      constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      122 a    121 c    110 g    178 t     10 others
ORIGIN


alignment_scores:
            Quality:  532.00            Length:    172
              Ratio:  3.855               Gaps:     10
 Percent Similarity:  80.233   Percent Identity: 71.512

alignment_block:
 US-10-031-904-8 x T83269   ..

 Align seg 1/1  to: T83269  from: 1  to: 541

       16 PheProGlyLeuLeuLeuAlaAlaLeuValLeuLeuLeuSerSerPheSe 32
          ||||||||||||||||||||||||||||||||||||||||||||||||||
        5 TTTCCTGGGTTGCTTCTGGCGGCCCTGGTGTTGCTGCTGTCCTCCTTCTC 54

       32 rAspGlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnL 49
          |||||||||||||||||||||||||||||||||||||||||||||||||||
       55 CGATCAATGCAATGTCCCGGAATGGCTTCCATTTGCCAGGCCTACCAACC 104

       49 euThrAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCys 65
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      105 TAACTGATGACTTTGAGTTTCCCATTGGGACATATCTGAACTATGAATGC 154

       66 ArgProGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSe 82
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      155 CGCCCTGGTTATTCCGGAAGACCGTTTTCTATCATCTGCCTAAAAAACTC 204

       82 rValTrpThrSerAlaLysAspLysCysLysArg.LysSerCysArgAsn 98
          |||||||||||||||||||||||||||||||| ||||||||||||||||
      205 AGTCTGGACAAGTGCTAAGGACAAGTGCAAACGTNAAATCATGTCGTAAT 254

       99 ProProAspProValAsn.GlyMetAlaHisVal..IleLysAspIleGl 114
          |||||||||||||||||| |||||||||||||||  :::||||||||||||
      255 CCTCCAGATCCTGTGAATGGGCATGGCACATGTGGATTCAAAGACATCCA 304

      114 nPheGly...SerGlnIleLysTyrSerCys.ProLysGly......Tyr 127
          ||||         ::::::||||||||||||| |||||||||
      305 GTTTCGGATTCCCAAATTTAAATATTCTTGTTCCTAAAGGGNTACCGANT 354

      128 ArgLeuIleGlySerSerSerAlaThrCysIleIleSerGlyAsnThrVa 144
          :::   :::::::||||||         ::::::||||||||||:::
      355 CATTGGGTTTCCTCGTCTTGCCACATGGCATTCATTTCAGGGCAACATTG 404

      144 lIleTrpAspAsn......LysThrProValCys...........AspS 155
            :::   :::      ||||||||||||                 :
      405 TTCATTTGGGGGGTTATTAAAAACACCCGTTTTGTGGACAGTNGAGTTTG 454

       155 erGluLeuLysTyrAlaPheLeuPheLeuLeuProIle.HisSerAsnPh 171
           ::   :::  :::   |||   |||   |||||||| |||      ||
       455 GAANTATGGCNTTCCCATTTTCCTTT...TTACCNATTACATTTTTATTT 501

       171 eSerLeu 173
           |:::|||
       502 TTTNCTC 508

seq_name: gb_est2:H73873

seq_documentation_block:
LOCUS       H73873                  440 bp    mRNA    linear   EST 31-OCT-1995
DEFINITION  ys14d08.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:214767 5' similar to gb:Y00816_cds1 COMPLEMENT RECEPTOR TYPE
            1 PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   H73873
VERSION     H73873.1  GI:1046874
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 440)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 706
            High quality sequence stops: 275
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 706   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 275.
FEATURES              Location/Qualifiers
     source           1. .440
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:214767"
                      /clone_lib="Soares fetal liver spleen 1NFLS"
                      /sex="male"
                      /dev_stage="20 week-post conception fetus"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                      with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                      1st strand cDNA was primed with a Pac I - oligo(dT) primer
                      [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Pac I and cloned into the Pac I
                      and Eco RI sites of the modified pT7T3 vector.  Library
                      went through one round of normalization. Library
                      constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       83 a    128 c    100 g    125 t      4 others
ORIGIN


alignment_scores:
            Quality:  497.00            Length:    118
              Ratio:  4.560               Gaps:      7
 Percent Similarity:  92.373   Percent Identity: 88.983
```

```
alignment_block:
 US-10-031-904-8 x H73873   ..

 Align seg 1/1  to: H73873  from: 1  to: 440

       4 ProValArgLeuGluArgProPheProSerArgArgPheProGlyLeuLe 20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
      93 CCCGTCCGTCTCGAGCGTCCCTTTCCTTCCCGGCGCTTTCCTGGGTTGCT 142

      20 uLeuAlaAlaLeuValLeuLeuLeuSerSerPheSerAspGlnCysAsnV 37
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     143 TCTGGCGGCCCTGGTGTTGCTGCTGTCCTCCTTCTCCGATCAATGCAATG 192

      37 alProGluTrpLeuProPheAlaArgProThrAsnLeuThrAspAspPhe 53
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
     193 TCCCGGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGACTTT 242

      54 GluPheProIleGlyThrTyrLeuAsnTyrGluCysArgPro.GlyTyrS 70
         |||||||||||||||||||||||||||||||||||||||| |||||||
     243 GAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGGTTATT 292

      70 erGlyArgProPheSerIleIleCysLeuLysAsnSerValTrp.ThrSe 86
         |||||||||||||||||||||||||||||||||||||||||| |||||
     293 CCGGAAGACCGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGGACAAG 342

      86 rAlaLysAspLysCysLysArg.LysSerCys.ArgAsnPro.ProAspP 102
         |||||||||||||||||||| |||||||| |||||||| |||||||
     343 TGCTAAGGACAAGTGCAAACGTTAAATCATGTTCGTAATCCTTCCAGATC 392

     102 roValAsn.GlyMetAlaHisValIle......LysAspIle 113
         |||||::: |||   :::|||:::::      |||||||||
     393 CTGTGGATTGGGCATNGGCACATTGTNGATCAAAAGGACATT 434

seq_name: gb_est1:AA212152

seq_documentation_block:
LOCUS       AA212152                601 bp    mRNA    linear   EST 31-JAN-1997
DEFINITION  mu80g07.r1 Stratagene mouse melanoma (#937312) Mus musculus cDNA
            clone IMAGE:651900 5' similar to gb:M23529 Mus musculus complement
            receptor (MOUSE);, mRNA sequence.
ACCESSION   AA212152
VERSION     AA212152.1  GI:1810867
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 601)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:397748
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 280.
FEATURES             Location/Qualifiers
     source          1. .601
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:651900"
                     /clone_lib="Stratagene mouse melanoma (#937312)"
                     /tissue_type="melanoma"
                     /dev_stage="M2 cells"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                     dT. From M2 cells, a highly metastatic derivative of the
                     K-1735 (mouse) melanoma. Average insert size: 1.0 kb;
                     Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG
                     3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT      149 a    137 c    140 g    175 t
ORIGIN


alignment_scores:
                Quality:  420.50              Length:    143
                  Ratio:  3.754                 Gaps:      1
 Percent Similarity:  78.322   Percent Identity: 55.245

alignment_block:
 US-10-031-904-8 x AA212152   ..

 Align seg 1/1  to: AA212152  from: 1  to: 601

      13 SerArgArgPheProGlyLeuLeuLeuAlaAlaLeuValLeuLeuLeuSe 29
         |||   |||      |||:::   |||   :::|||:::|||:::|||
      69 AGCCTTCGGCGGGGAGGAGTCAAGCTAGAAGTTTGCTGCTGTTCTTGCT 118

      29 rSerPheSer...AspGlnCysAsnValProGluTrpLeuProPheAlaA 45
           |||:::      :::|||   :::|||:::   ||||||   |||:
     119 GCCATTTACTTTGGGTCACTGCCCAGCCCCATCACAGCTTCCTTCTGCCA 168

      45 rgProThrAsnLeuThrAspAspPheGluPheProIleGlyThrTyrLeu 61
         ::|||   |||||||||||||:::      ||||||||||||||||||||
     169 AACCTATAAATCTAACTGATGAATCCATGTTTCCCATTGGAACATATTTG 218

      62 AsnTyrGluCysArgProGlyTyrSerGlyArgProPheSerIleIleCy 78
         |||||||||   |||||||||      |||   |||||||||   ||
     219 TTGTATGAATGTCTCCCAGGATATATCAAGAGGCAGTTCTCTATCACCTG 268

      78 sLeuLysAsnSerValTrpThrSerAlaLysAspLysCysLysArgLysS 95
         |   ::::::|||:::||||||||||||:::|||||||||   ||||||:
     269 CAAACAAGACTCAACCTGGACGAGTGCTGAAGATAAGTGTATACGAAAAC 318

      95 erCysArgAsnProProAspProValAsnGlyMetAlaHisValIleLys 111
         ::|||:::::|||   ||||||   ||||||:::::|||||
     319 AATGTAAAACTCCTTCAGATCCTGAGAATGGCTTGGTACATGTACACACA 368

     112 AspIleGlnPheGlySerGlnIleLysTyrSerCysProLysGlyTyrAr 128
         ||||||||||||||||:::|||:::|||:::|||   :::|||||||||
     369 GGCATTCAGTTTGGATCCCGTATTAATTATACTTGTAATCAAGGATACCG 418

     128 gLeuIleGlySerSerSerAlaThrCysIleIleSerGlyAsnThrValI 145
         ||||||||||||||||||||:::|||:::|||:::   ::::::|||
     419 CCTCATTGGTTCCTCCTCTGCTGTATGTGTCATCACTGATCAAAGTGTTG 468

     145 leTrpAspAsnLysThrProValCysAsp 154
         ||||||:::::::::|||:::|||:::
     469 ATTGGGATACTGAGGCACCCATTTGTGAG 497

seq_name: gb_est2:BG077250

seq_documentation_block:
LOCUS       BG077250                577 bp    mRNA    linear   EST 26-JAN-2001
DEFINITION  H3013A12-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
            H3013A12 5', mRNA sequence.
ACCESSION   BG077250
VERSION     BG077250.1  GI:12559818
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 577)
```

```
  AUTHORS   Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka
            ,T.S., Carter,M.G. and Ko,M.S.H.
  TITLE     Verification and initial annotation of NIA mouse 15K cDNA clone set
  JOURNAL   Unpublished (2001)
COMMENT     Other_ESTs: H3013A12-3
            Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            This clone set has been freely distributed to the community. Please
            visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details.
            Plate: H3013  row: A  column: 12
            Seq primer: -21M13 Reverse
            High quality sequence stop: 577
            POLYA=NO.
FEATURES             Location/Qualifiers
     source          1. .577
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="niaEST:H3013A12-5"
                     /db_xref="taxon:10090"
                     /clone="H3013A12"
                     /clone_lib="NIA Mouse 15K cDNA Clone Set"
                     /sex="Clones arrayed from a variety of cDNA libraries"
                     /dev_stage="Clones arrayed from a variety of cDNA
                     libraries"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
                     clone is among a rearrayed set of 15,247 clones from 11
                     embryo  cDNA libraries (including preimplantation stage
                     embryos from  unfertilized egg to blastocyst, embryonic
                     part of E7.5 embryos,  extraembryonic part of E7.5 embryos
                     , and E12.5 female mesonephros/gonad)  and one newborn
                     ovary cDNA library. Average insert size 1.5 kb. All
                     source libraries are cloned unidirectionally with Oligo(dT
                     )-Not primers.  References include: (1) Genome-wide
                     expression profiling of mid-gestation  placenta and embryo
                     using a 15,000 mouse developmental cDNA microarray,  2000,
                     Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2)
                     Large-scale cDNA  analysis reveals phased gene expression
                     patterns during preimplantation  mouse develolpment, 2000,
                     Development, 127: 1737-1749; (3) Genome-wide  mapping of
                     unselected transcripts from extraembryonic tissue of
                     7.5-day  mouse embryos reveals enrichment in the t-complex
                     and under-representation  on the X chromosome, 1998, Hum
                     Mol Genet 7: 1967-1978."
BASE COUNT      141 a    132 c    132 g    172 t
ORIGIN


alignment_scores:
            Quality:  418.50            Length:    138
              Ratio:  3.805               Gaps:      1
 Percent Similarity:  79.710   Percent Identity: 55.797

alignment_block:
 US-10-031-904-8 x BG077250   ..

 Align seg 1/1  to: BG077250  from: 1  to: 577

      18 GlyLeuLeuLeuAlaAlaLeuValLeuLeuLeuSerSerPheSer...As 33
         |||:::   |||   :::|||:::|||:::|||      |||:::
     122 GGAGTCAAGCTAGAAGTTTGCTGCTGTTCTTGCTGCCATTTACTTTGGG 171

      33 pGlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuT 50
          :::|||   :::|||:::   ||||||   |||:::|||   |||||||
     172 TCACTGCCCAGCCCCATCACAGCTTCCTTCTGCCAAACCTATAAATCTAA 221

      50 hrAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArg 66
         |||||:::      ||||||||||||||||||||   |||||||||
     222 CTGATGAATCCATGTTTCCCATTGGAACATATTTGTTGTATGAATGTCTC 271

      67 ProGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVa 83
         |||||||||      |||   |||||||||   |||   ::::::|||::
     272 CCAGGATATATCAAGAGGCAGTTCTCTATCACCTGCAAACAAGACTCAAC 321

      83 lTrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProP 100
         :|||||||||||||:::|||||||||   ||||||:::|||:::::::|||
     322 CTGGACGAGTGCTGAAGATAAGTGTATACGAAAACAATGTAAAACTCCTT 371

     100 roAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGly 116
           ||||||   ||||||:::::||||||          ||||||||||||
     372 CAGATCCTGAGAATGGCTTGGTACATGTACACACAGGCATTCAGTTTGGA 421

     117 SerGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSe 133
         |||:::|||:::|||:::|||   :::|||||||||||||||||||||||
     422 TCCCGTATTAATTATACTTGTAATCAAGGATACCGCCTCATTGGTTCCTC 471

     133 rSerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysT 150
         |||||||:::|||:::|||:::   :::::::|||   ||||||:::::::
     472 CTCTGCTGTATGTGTCATCACTGATCAAAGTGTTGATTGGGATACTGAGG 521

     150 hrProValCysAsp 154
         ::|||:::|||:::
     522 CACCTATTTGTGAG 535

seq_name: gb_est2:BI697900

seq_documentation_block:
LOCUS       BI697900               658 bp    mRNA    linear   EST 18-SEP-2001
DEFINITION  603344033F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5371665 5',
            mRNA sequence.
ACCESSION   BI697900
VERSION     BI697900.1  GI:15660529
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 658)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11946  row: a  column: 10
            High quality sequence stop: 658.
FEATURES             Location/Qualifiers
     source          1. .658
                     /organism="Mus musculus"
                     /strain="FVB/N-3"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5371665"
                     /clone_lib="NCI_CGAP_Mam2"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT      163 a    146 c    158 g    191 t
ORIGIN


alignment_scores:
            Quality:  418.50            Length:    138
              Ratio:  3.805               Gaps:      1
```

```
 Percent Similarity:  79.710   Percent Identity: 55.797

alignment_block:
 US-10-031-904-8 x BI697900   ..

 Align seg 1/1  to: BI697900  from: 1  to: 658

      18 GlyLeuLeuLeuAlaAlaLeuValLeuLeuLeuSerSerPheSer...As 33
         |||:::   |||   :::|||:::|||:::|||      |||:::
     102 GGAGTCAAGCTAGAAGTTTTGCTGCTGTTCTTGCTGCCATTTACTTTGGG 151

      33 pGlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuT 50
          :::|||   :::|||:::   ||||||   |||:::|||   |||||||
     152 TCACTGCCCAGCCCCATCACAGCTTCCTTCTGCCAAACCTATAAATCTAA 201

      50 hrAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArg 66
         |||||:::      ||||||||||||||||||||   |||||||||
     202 CTGATGAATCCATGTTTCCCATTGGAACATATTTGTTGTATGAATGTCTC 251

      67 ProGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVa 83
         |||||||||      |||   |||||||||   |||   ::::::|||::
     252 CCAGGATATATCAAGAGGCAGTTCTCTATCACCTGCAAACAAGACTCAAC 301

      83 lTrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProP 100
         :||||||||||||:::|||||||||   ||||||:::|||::::::|||
     302 CTGGACGAGTGCTGAAGATAAGTGTATACGAAAACAATGTAAAACTCCTT 351

     100 roAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGly 116
           ||||||   ||||||:::::||||||         ||||||||||||
     352 CAGATCCTGAGAATGGCTTGGTACATGTACACACAGGCATTCAGTTTGGA 401

     117 SerGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSe 133
         |||:::|||:::|||:::|||   :::||||||||||||||||||||||
     402 TCCCGTATTAATTATACTTGTAATCAAGGATACCGCCTCATTGGTTCCTC 451

     133 rSerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysT 150
         |||||||:::|||:::|||:::   ::::::|||   ||||||:::::::
     452 CTCTGCTGTATGTGTCATCACTGATCAAAGTGTTGATTGGGATACTGAGG 501

     150 hrProValCysAsp 154
         ::|||:::|||:::
     502 CACCTATTTGTGAG 515

seq_name: gb_est2:BI455761

seq_documentation_block:
LOCUS       BI455761                795 bp    mRNA    linear   EST 21-AUG-2001
DEFINITION  603171769F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5251289 5',
            mRNA sequence.
ACCESSION   BI455761
VERSION     BI455761.1  GI:15246417
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 795)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11634  row: i  column: 18
            High quality sequence stop: 795.
FEATURES             Location/Qualifiers
     source          1. .795
                     /organism="Mus musculus"
                     /strain="C57/B6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5251289"
                     /clone_lib="NCI_CGAP_Mam5"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="7 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Robin Humphreys,
                     NIH"
BASE COUNT      191 a    192 c    186 g    226 t
ORIGIN


alignment_scores:
           Quality:  418.50             Length:    138
             Ratio:  3.805                Gaps:      1
 Percent Similarity:  79.710   Percent Identity: 55.797

alignment_block:
 US-10-031-904-8 x BI455761   ..

 Align seg 1/1  to: BI455761  from: 1  to: 795

      18 GlyLeuLeuLeuAlaAlaLeuValLeuLeuLeuSerSerPheSer...As 33
         |||:::   |||   :::|||:::|||:::|||      |||:::
     140 GGAGTCAAGCTAGAAGTTTTGCTGCTGTTCTTGCTGCCATTTACTTTGGG 189

      33 pGlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuT 50
          :::|||   :::|||:::   ||||||   |||:::|||   |||||||
     190 TCACTGCCCAGCCCCATCACAGCTTCCTTCTGCCAAACCTATAAATCTAA 239

      50 hrAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArg 66
         |||||:::      ||||||||||||||||||||   |||||||||
     240 CTGATGAATCCATGTTTCCCATTGGAACATATTTGTTGTATGAATGTCTC 289

      67 ProGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVa 83
         |||||||||      |||   |||||||||   |||   ::::::|||::
     290 CCAGGATATATCAAGAGGCAGTTCTCTATCACCTGCAAACAAGACTCAAC 339

      83 lTrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProP 100
         :||||||||||||:::|||||||||   ||||||:::|||::::::|||
     340 CTGGACGAGTGCTGAAGATAAGTGTATACGAAAACAATGTAAAACTCCTT 389

     100 roAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGly 116
           ||||||   ||||||:::::||||||         ||||||||||||
     390 CAGATCCTGAGAATGGCTTGGTACATGTACACACAGGCATTCAGTTTGGA 439

     117 SerGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSe 133
         |||:::|||:::|||:::|||   :::||||||||||||||||||||||
     440 TCCCGTATTAATTATACTTGTAATCAAGGATACCGCCTCATTGGTTCCTC 489

     133 rSerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysT 150
         |||||||:::|||:::|||:::   :::::::||   ||||||:::::::
     490 CTCTGCTGTATGTGTCATCACTGATCAAAGTGTTGATTGGGATACTGAGG 539

     150 hrProValCysAsp 154
         ::|||:::|||:::
     540 CACCTATTTGTGAG 553

seq_name: gb_est2:BI078162

seq_documentation_block:
LOCUS       BI078162                807 bp    mRNA    linear   EST 20-JUN-2001
DEFINITION  602872480F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5004523 5',
            mRNA sequence.
ACCESSION   BI078162
VERSION     BI078162.1  GI:14496492
KEYWORDS    EST.
SOURCE      house mouse.
```

```
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 807)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11042  row: o  column: 20
            High quality sequence stop: 652.
FEATURES             Location/Qualifiers
     source          1. .807
                     /organism="Mus musculus"
                     /strain="FVB/N-3"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5004523"
                     /clone_lib="NCI_CGAP_Mam2"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT      202 a    188 c    190 g    227 t
ORIGIN


alignment_scores:
            Quality:  418.50             Length:    138
              Ratio:  3.805                Gaps:      1
 Percent Similarity:  79.710   Percent Identity: 55.797

alignment_block:
 US-10-031-904-8 x BI078162   ..

 Align seg 1/1  to: BI078162  from: 1  to: 807

      18 GlyLeuLeuLeuAlaAlaLeuValLeuLeuLeuSerSerPheSer...As 33
         |||:::   |||   :::|||:::|||:::|||      |||:::
      85 GGAGTCAAGCTAGAAGTTTTGCTGCTGTTCTTGCTGCCATTTACTTTGGG 134

      33 pGlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuT 50
          :::|||   :::|||:::   ||||||   |||:::|||   |||||||
     135 TCACTGCCCAGCCCCATCACAGCTTCCTTCTGCCAAACCTATAAATCTAA 184

      50 hrAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArg 66
         |||||:::      ||||||||||||||||||||   |||||||||
     185 CTGATGAATCCATGTTTCCCATTGGAACATATTTGTTGTATGAATGTCTC 234

      67 ProGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVa 83
         |||||||||      |||   |||||||||   |||   ::::::|||::
     235 CCAGGATATATCAAGAGGCAGTTCTCTATCACCTGCAAACAAGACTCAAC 284

      83 lTrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProP 100
         :||||||||||||:::|||||||||   |||||||:::|||:::::|||
     285 CTGGACGAGTGCTGAAGATAAGTGTATACGAAAACAATGTAAAACTCCTT 334

     100 roAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGly 116
           ||||||   ||||||::::::||||||         ||||||||||||
     335 CAGATCCTGAGAATGGCTTGGTACATGTACACACAGGCATTCAGTTTGGA 384

     117 SerGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSe 133
         |||:::|||:::|||:::|||   :::|||||||||||||||||||||||
     385 TCCCGTATTAATTATACTTGTAATCAAGGATACCGCCTCATTGGTTCCTC 434

     133 rSerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysT 150
         |||||||:::|||:::|||:::   :::::::|||   ||||||:::::::
     435 CTCTGCTGTATGTGTCATCACTGATCAAAGTGTTGATTGGGATACTGAGG 484

     150 hrProValCysAsp 154
         ::|||:::|||:::
     485 CACCTATTTGTGAG 498
```

```
OM of: US-10-031-904-8 to: Issued_Patents_NA:*   out_format : pfs

Date: Oct 9, 2002  6:38 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlh
 -Q=/cgn2_1/USPTO_spool/US10031904/runat_09102002_084256_23659/app_query.fasta_1.232
 -DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni
 -GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000
 -LOOPEXT=0.000 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000
 -XGAPEXT=0.500 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000
 -YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -START=1
 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200
 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
 -OUTFMT=pfs -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 -USER=US10031904_@CGN1_1_68 -NCPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-10-031-904-8
Query length: 174
Database: Issued_Patents_NA:*
Database sequences: 383533
Database length: 122816752
Search time (sec): 41.920000


score_list:
Sequence           Strd Orig    %Score   EScore Len  ! Documentation  ..
/cgn2_6/ptodata/2/ina/backfiles1.seq:5256642-1 +   649.50 1363.82  2.8e-68   6951  ! Pate
/cgn2_6/ptodata/2/ina/backfiles1.seq:5472939-1 +   649.50 1363.82  2.8e-68   6951  ! Pate
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-769-967A-33 +   635.00 1360.80  4.1e-68   605  !
/cgn2_6/ptodata/2/ina/backfiles1.seq:5256642-3 +   442.00  914.45  3.0e-43   5420  ! Pate
/cgn2_6/ptodata/2/ina/backfiles1.seq:5472939-3 +   442.00  914.45  3.0e-43   5420  ! Pate
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-139-195-1 +   291.00  600.02  9.8e-26   1545  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-310-416A-11 +   289.00  601.95  7.7e-26   903  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-888-171-11 +   289.00  601.95  7.7e-26   903  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-793-418-1 +   289.00  599.28  1.1e-25   1134  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-793-418-3 +   289.00  599.28  1.1e-25   1134  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-528-057-41 +   289.00  598.17  1.2e-25   1247  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-528-057-45 +   289.00  597.65  1.3e-25   1304  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-458-084-3 +   289.00  595.78  1.7e-25   1530  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-205-508-3 +   289.00  595.78  1.7e-25   1530  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-482-148-8 +   289.00  595.78  1.7e-25   1530  !
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US95-02944-8 +   289.00  595.78  1.7e-25   1530
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US95-02945-3 +   289.00  595.78  1.7e-25   1530
/cgn2_6/ptodata/2/ina/backfiles1.seq:5514787-1 +   289.00  595.65  1.7e-25   1546  ! Pate
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-528-057-1 +   289.00  594.83  1.9e-25   1659  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-475-460A-31 +   289.00  593.57  2.3e-25   1848  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-435-149-17 +   289.00  593.38  2.3e-25   1878  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-528-057-43 +   289.00  592.72  2.5e-25   1986  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-087-007-1 +   289.00  591.94  2.8e-25   2124  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-483-433-1 +   289.00  591.94  2.8e-25   2124  !
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US92-05920-1 +   289.00  591.94  2.8e-25   2124
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-793-418-2 +   287.00  594.92  1.9e-25   1134  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-793-418-4 +   287.00  594.92  1.9e-25   1134  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-811-048-9 +   272.50  556.02  2.8e-23   2116  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-811-048-10 +   272.50  555.38  3.0e-23   2234  !
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-981-234B-1 +   271.50  558.97  1.9e-23   1365  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-310-416A-19 +   265.00  550.21  5.9e-23   860  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-888-171-19 +   265.00  550.21  5.9e-23   860  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-310-416A-12 +   265.00  548.81  7.0e-23   969  !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-888-171-12 +   265.00  548.81  7.0e-23   969  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-458-084-1 +   265.00  539.78  2.2e-22   2096  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-205-508-1 +   265.00  539.78  2.2e-22   2096  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-278-630A-10 +   265.00  539.78  2.2e-22   2096  !
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US95-02945-1 +   265.00  539.78  2.2e-22   2096
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-087-007-2 +   265.00  536.19  3.5e-22   2847  !
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-483-433-2 +   265.00  536.19  3.5e-22   2847  !
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US92-05920-2 +   265.00  536.19  3.5e-22   28
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-793-418-5 +   245.50  506.60  1.6e-20   945  !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-278-630A-13 +   221.50  446.83  3.4e-17   1787
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-278-630A-12 +   221.50  445.63  3.9e-17   1980
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-356-361-3 +   178.00  385.66  8.6e-14   101  !

seq_name: /cgn2_6/ptodata/2/ina/backfiles1.seq:5256642-1

seq_documentation_block:
;Patent No. 5256642
;     APPLICANT: FEARON, DOUGLAS T.;KLICKSTEIN, LLOYD B.;WONG,
;WINNIE W.;CARSON, GERALD R.;CONCINO, MICHAEL F.;IP, STEPHEN
;H.;MAKRIDES, SAVVAS;MARSH, HENRY C. JR.
;     TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
;RECEPTOR 1 (CR1) AND A THROMBOLTIC AGENT, AND THE METHODS OF
;USE THEREOF
;     NUMBER OF SEQUENCES: 30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/588,128
;       FILING DATE: 24-SEP-1990
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 412,745
;       FILING DATE: 26-SEP-1989
;       APPLICATION NUMBER: 332,865
;       FILING DATE: 03-APR-1989
;       APPLICATION NUMBER: 176,532
;       FILING DATE: 01-APR-1988
;SEQ ID NO:1:
;       LENGTH: 6951
5256642-1


alignment_scores:
            Quality:  649.50              Length:    148
              Ratio:  4.883                 Gaps:      2
 Percent Similarity:  89.865   Percent Identity: 81.757

alignment_block:
 US-10-031-904-8 x 5256642-1   ..

 Align seg 1/1  to: 5256642-1  from: 1  to: 6951

      10 ProPheProSerArgArgPhe......ProGlyLeuLeuLeuAlaAlaLe 24
         |||   |||:::     |||        |||   |||||||||:::::
      70 CCGGCGCCCGGTCTCCCCTTCTGCTGCGGAGGATCCCTGCTGGCGGTTGT 119

      24 uValLeuLeu...LeuSerSerPheSerAspGlnCysAsnValProGluT 40
         :|||||||||   |||          |||||||||:::|||||||
     120 GGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAGAAT 169

      40 rpLeuProPheAlaArgProThrAsnLeuThrAspAspPheGluPhePro 56
         ||||||||||||||||||||||||||||||||:::|||||||||||
     170 GGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCCC 219

      57 IleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgPr 73
         ||||||||||||||||||||||||||||||||||||||||||||||||
     220 ATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGACC 269

      73 oPheSerIleIleCysLeuLysAsnSerValTrpThrSerAlaLysAspL 90
         ||||||||||||||||||||||||||||||||:::|||||||||:
     270 GTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACTGGTGCTAAGGACA 319

      90 ysCysLysArgLysSerCysArgAsnProProAspProValAsnGlyMet 106
         ::|||:::|||||||||||||||||||||||||||||||||||||||||
     320 GGTGCAGACGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATG 369

     107 AlaHisValIleLysAspIleGlnPheGlySerGlnIleLysTyrSerCy 123
         :::||||||||||||   ||||||||||||||||||||||||||||||||
     370 GTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAAATTAAATATTCTTG 419

     123 sProLysGlyTyrArgLeuIleGlySerSerSerAlaThrCysIleIleS 140
         |   ||||||||||||||||||||||||||||||||||||||||||||||
     420 TACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCATCT 469
```

```
     140 erGlyAsnThrValIleTrpAspAsnLysThrProValCysAsp 154
         |||||:::||||||||||||||||||||:::||||||:::||||||
     470 CAGGTGATACTGTCATTTGGGATAATGAAACACCTATTTGTGAC 513

seq_name: /cgn2_6/ptodata/2/ina/backfiles1.seq:5472939-1

seq_documentation_block:
;Patent No. 5472939
;    APPLICANT: FEARON, DOUGLAS T.;KLICKSTEIN, LLOYD B.;WONG,
;WINNIE W.;CARSON, GERALD R.;CONCINO, MICHAEL F.;IP, STEPHEN
;H.;MAKRIDES, SAVVAS;MARSH, HENRY C. JR.
;    TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
;MEDIATED DISORDERS
;    NUMBER OF SEQUENCES: 30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/138,825
;      FILING DATE: 19-OCT-1993
;    PRIOR APPLICATION DATA:
       APPLICATION NUMBER: 588,128
       FILING DATE: 24-SEP-1990
       APPLICATION NUMBER: 412,745
;      FILING DATE: 26-SEP-1989
;      APPLICATION NUMBER: 332,865
;      FILING DATE: 03-APR-1989
;      APPLICATION NUMBER: 176,532
;      FILING DATE: 01-APR-1988
;SEQ ID NO:1:
;      LENGTH: 6951
5472939-1


alignment_scores:
            Quality:  649.50            Length:    148
              Ratio:  4.883              Gaps:      2
 Percent Similarity:  89.865   Percent Identity: 81.757

alignment_block:
 US-10-031-904-8 x 5472939-1   ..

 Align seg 1/1  to: 5472939-1  from: 1  to: 6951

      10 ProPheProSerArgArgPhe......ProGlyLeuLeuLeuAlaAlaLe 24
         |||   |||:::      |||         |||   |||||||||:::::
      70 CCGGCGCCCGGTCTCCCCTTCTGCTGCGGAGGATCCCTGCTGGCGGTTGT 119

      24 uValLeuLeu...LeuSerSerPheSerAspGlnCysAsnValProGluT 40
         :|||||||||   |||               ||||||||||:::|||||||
     120 GGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAGAAT 169

      40 rpLeuProPheAlaArgProThrAsnLeuThrAspAspPheGluPhePro 56
         ||||||||||||||||||||||||||||||||||||:::|||||||||||
     170 GGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCCC 219

      57 IleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgPr 73
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     220 ATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGACC 269

      73 oPheSerIleIleCysLeuLysAsnSerValTrpThrSerAlaLysAspL 90
         |||||||||||||||||||||||||||||||||||:::|||||||||:
     270 GTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACTGGTGCTAAGGACA 319

      90 ysCysLysArgLysSerCysArgAsnProProAspProValAsnGlyMet 106
         ::|||:::||||||||||||||||||||||||||||||||||||||||||
     320 GGTGCAGACGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATG 369

     107 AlaHisValIleLysAspIleGlnPheGlySerGlnIleLysTyrSerCy 123
         :::||||||||||||   ||||||||||||||||||||||||||||||||
     370 GTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAAATTAAATATTCTTG 419

     123 sProLysGlyTyrArgLeuIleGlySerSerSerAlaThrCysIleIleS 140
         |   ||||||||||||||||||||||||||||||||||||||||||||||
     420 TACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCATCT 469
```

```
     140 erGlyAsnThrValIleTrpAspAsnLysThrProValCysAsp 154
         |||||:::||||||||||||||||||||:::||||||:::||||||
     470 CAGGTGATACTGTCATTTGGGATAATGAAACACCTATTTGTGAC 513

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-769-967A-33

seq_documentation_block:
; Sequence 33, Application US/08769967A
; Patent No. 5859223
;  GENERAL INFORMATION:
;    APPLICANT:  Mossakowska, Danuta E.I.
;    APPLICANT:  Smith, Richard A.G.
;    APPLICANT:  Dodd, Ian
;    APPLICANT:  Freeman, Anne Mary
;    TITLE OF INVENTION:  Soluble CR1 Derivatives
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporate Intellectual Property
;      STREET:  P.O. Box 1539
;      CITY:  King of Prussia
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19406
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/769,967A
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/440,569
;      FILING DATE:  15-May-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  King, William T.
;      REGISTRATION NUMBER:  30,954
;      REFERENCE/DOCKET NUMBER:  P30423C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610) 270-5364
;      TELEFAX:  (610) 270-5090
;  INFORMATION FOR SEQ ID NO:  33:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  605 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-769-967A-33


alignment_scores:
            Quality:  635.00            Length:    121
              Ratio:  5.336              Gaps:      0
 Percent Similarity:  98.347   Percent Identity: 90.909

alignment_block:
 US-10-031-904-8 x US-08-769-967A-33   ..

 Align seg 1/1  to: US-08-769-967A-33  from: 1  to: 605

      34 GlnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuTh 50
         ||||||||||:::||||||||||||||||||||||||||||||||||||||
       7 CAGTGCAACGCTCCGGAATGGCTGCCGTTCGCGCGCCCGACCAACCTGAC 56

      50 rAspAspPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgP 67
         ||||:::||||||||||||||||||||||||||||||||||||||||||||
      57 TGATGAATTTGAGTTCCCGATCGGTACCTACCTGAACTACGAATGCCGCC 106

      67 roGlyTyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerVal 83
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
    107 CGGGTTATAGCGGCCGCCCGTTTTCTATCATCTGCCTGAAAAACTCTGTC 156

     84 TrpThrSerAlaLysAspLysCysLysArgLysSerCysArgAsnProPr 100
        ||||||:::|||||||||:::|||:::||||||||||||||||||||||||
    157 TGGACTGGTGCTAAGGACCGTTGCCGACGTAAATCTTGTCGTAATCCGCC 206

    100 oAspProValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlyS 117
        ||||||||||||||||||:::||||||||||||   ||||||||||||
    207 AGATCCGGTTAACGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGTT 256

    117 erGlnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSer 133
        ||||||||||||||||||   ||||||||||||||||||||||||||||
    257 CCCAAATTAAATATTCTTGTACTAAAGGTTACCGTCTGATTGGTTCCTCC 306

    134 SerAlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysTh 150
        ||||||||||||||||||||||||:::||||||||||||||||||:::||
    307 AGCGCTACATGCATCATCTCTGGTGATACTGTCATTTGGGATAATGAAAC 356

    150 rProValCysAsp 154
        ||||:::|||||||
    357 ACCGATTTGTGAC 369
```

seq_name: /cgn2_6/ptodata/2/ina/backfiles1.seq:5256642-3

seq_documentation_block:
;Patent No. 5256642
; APPLICANT: FEARON, DOUGLAS T.;KLICKSTEIN, LLOYD B.;WONG,
;WINNIE W.;CARSON, GERALD R.;CONCINO, MICHAEL F.;IP, STEPHEN
;H.;MAKRIDES, SAVVAS;MARSH, HENRY C. JR.
; TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
;RECEPTOR 1 (CR1) AND A THROMBOLTIC AGENT, AND THE METHODS OF
;USE THEREOF
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/588,128
;   FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 412,745
;   FILING DATE: 26-SEP-1989
;   APPLICATION NUMBER: 332,865
;   FILING DATE: 03-APR-1989
;   APPLICATION NUMBER: 176,532
;   FILING DATE: 01-APR-1988
;SEQ ID NO:3:
;   LENGTH: 5420
5256642-3

gnment_scores:
Quality: 442.00 Length: 127
Ratio: 4.250 Gaps: 0
Percent Similarity: 81.890 Percent Identity: 59.055

alignment_block:
US-10-031-904-8 x 5256642-3 ..

Align seg 1/1 to: 5256642-3 from: 1 to: 5420

```
     28 LeuSerSerPheSerAspGlnCysAsnValProGluTrpLeuProPheAl 44
        ||||||      :::   :::|||:::::::||||||   :::||||||||
   2661 CTTTCTGTTCGTGCTGGTCACTGTAAAACCCCAGAGCAGTTTCCATTTGC 2710

     44 aArgProThrAsnLeuThrAspAspPheGluPheProIleGlyThrTyrL 61
        |   ||||||         :::||||||||||||||||:::||||||   |
   2711 CAGTCCTACGATCCCAATTAATGACTTTGAGTTTCCAGTCGGGACATCTT 2760

     61 euAsnTyrGluCysArgProGlyTyrSerGlyArgProPheSerIleIle 77
        ||||||||||||||||||||||||   |||:::   |||||||||
   2761 TGAATTATGAATGCCGTCCTGGGTATTTTGGGAAAATGTTCTCTATCTCC 2810

     78 CysLeuLysAsnSerValTrpThrSerAlaLysAspLysCysLysArgLy 94
        ||||||:::|||   ||||||:::|||:::::|||:::|||:::||||||
   2811 TGCCTAGAAAACTTGGTCTGGTCAAGTGTTGAAGACAACTGTAGACGAAA 2860

     94 sSerCysArgAsnProProAspProValAsnGlyMetAlaHisValIleL 111
        |||||||      ||||||:::|||   ||||||||||:::|||:::
   2861 ATCATGTGGACCTCCACCAGAACCCTTCAATGGAATGGTGCATATAAACA 2910

    111 ysAspIleGlnPheGlySerGlnIleLysTyrSerCysProLysGlyTyr 127
          |||   ||||||||||||   ::::::||||||||||   :::|||:::
   2911 CAGATACACAGTTTGGATCAACAGTTAATTATTCTTGTAATGAAGGGTTT 2960

    128 ArgLeuIleGlySerSerSerAlaThrCysIleIleSerGlyAsnThrVa 144
        ||||||||||||||   |||:::||||||::::::||||||||||:::||
   2961 CGACTCATTGGTTCCCCATCTACTACTTGTCTCGTCTCAGGCAATAATGT 3010

    144 lIleTrpAspAsnLysThrProValCysAsp 154
        |   ||||||:::|||:::|||:::|||:::
   3011 CACATGGGATAAGAAGGCACCTATTTGTGAG 3041
```

seq_name: /cgn2_6/ptodata/2/ina/backfiles1.seq:5472939-3

seq_documentation_block:
;Patent No. 5472939
; APPLICANT: FEARON, DOUGLAS T.;KLICKSTEIN, LLOYD B.;WONG,
;WINNIE W.;CARSON, GERALD R.;CONCINO, MICHAEL F.;IP, STEPHEN
;H.;MAKRIDES, SAVVAS;MARSH, HENRY C. JR.
; TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
;MEDIATED DISORDERS
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/138,825
;   FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 588,128
;   FILING DATE: 24-SEP-1990
;   APPLICATION NUMBER: 412,745
;   FILING DATE: 26-SEP-1989
;   APPLICATION NUMBER: 332,865
;   FILING DATE: 03-APR-1989
;   APPLICATION NUMBER: 176,532
;   FILING DATE: 01-APR-1988
;SEQ ID NO:3:
;   LENGTH: 5420
5472939-3

alignment_scores:
Quality: 442.00 Length: 127
Ratio: 4.250 Gaps: 0
Percent Similarity: 81.890 Percent Identity: 59.055

alignment_block:
US-10-031-904-8 x 5472939-3 ..

Align seg 1/1 to: 5472939-3 from: 1 to: 5420

```
     28 LeuSerSerPheSerAspGlnCysAsnValProGluTrpLeuProPheAl 44
        ||||||      :::   :::|||:::::::||||||   :::||||||||
   2661 CTTTCTGTTCGTGCTGGTCACTGTAAAACCCCAGAGCAGTTTCCATTTGC 2710

     44 aArgProThrAsnLeuThrAspAspPheGluPheProIleGlyThrTyrL 61
        |   ||||||         :::||||||||||||||||:::||||||   |
   2711 CAGTCCTACGATCCCAATTAATGACTTTGAGTTTCCAGTCGGGACATCTT 2760

     61 euAsnTyrGluCysArgProGlyTyrSerGlyArgProPheSerIleIle 77
        ||||||||||||||||||||||||   |||:::   |||||||||
   2761 TGAATTATGAATGCCGTCCTGGGTATTTTGGGAAAATGTTCTCTATCTCC 2810

     78 CysLeuLysAsnSerValTrpThrSerAlaLysAspLysCysLysArgLy 94
        ||||||:::|||   ||||||:::|||:::::|||:::|||:::||||||
   2811 TGCCTAGAAAACTTGGTCTGGTCAAGTGTTGAAGACAACTGTAGACGAAA 2860

     94 sSerCysArgAsnProProAspProValAsnGlyMetAlaHisValIleL 111
        |||||||-     ||||||:::|||   ||||||||||:::|||:::
   2861 ATCATGTGGACCTCCACCAGAACCCTTCAATGGAATGGTGCATATAAACA 2910
```

```
     111 ysAspIleGlnPheGlySerGlnIleLysTyrSerCysProLysGlyTyr 127
           |||   ||||||||||||   ::::::||||||||   :::|||:::
    2911 CAGATACACAGTTTGGATCAACAGTTAATTATTCTTGTAATGAAGGGTTT 2960

     128 ArgLeuIleGlySerSerSerAlaThrCysIleIleSerGlyAsnThrVa 144
         ||||||||||||||   |||:::||||||::::::|||||||||:::||
    2961 CGACTCATTGGTTCCCCATCTACTACTTGTCTCGTCTCAGGCAATAATGT 3010

     144 lIleTrpAspAsnLysThrProValCysAsp 154
         |   ||||||:::|||:::|||:::|||:::
    3011 CACATGGGATAAGAAGGCACCTATTTGTGAG 3041

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-139-195-1

seq_documentation_block:
; Sequence 1, Application US/08139195
; Patent No. 6218520
;  GENERAL INFORMATION:
;    APPLICANT:  Atkinson, John P.
;    TITLE OF INVENTION:  RECOMBINANTLY PRODUCED HUMAN MEMBRANE
;    TITLE OF INVENTION:  COFACTOR PROTEIN (MCP)
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  1100 Peachtree Street, Suite 2800
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-4530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/139,195
;      FILING DATE:  20-OCT-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/948,350
;      FILING DATE:  21-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/384,210
;      FILING DATE:  21-JUL-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  WU100CON
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404)-815-6508
;      TELEFAX:  (404)-815-6555
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1545 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Human t-Cell Line HSB2
US-08-139-195-1


alignment_scores:
            Quality:  291.00            Length:    168
              Ratio:  2.853               Gaps:      6
 Percent Similarity:  60.714   Percent Identity: 40.476

alignment_block:
 US-10-031-904-8 x US-08-139-195-1   ..
```

```
 Align seq 1/1  to: US-08-139-195-1  from: 1  to: 1545

       1 MetAlaProProValArgLeuGluArgProPheProSerArgArgPhePr 17
         |||   ||||||   |||   |||   ||||||||||||   ||||||||
      44 ATGGAGCCTCCCGGCCGCCGCGAGTGTCCCTTTCCTTCCTGGCGCTTTCC 93

      17 oGlyLeuLeuLeuAlaAlaLeuValLeuLeuLeuSerSerPheSerAspG 34
         |||||||||||||||||||:::||||||||||||   ||||||||||||
      94 TGGGTTGCTTCTGGCGGCCATGGTGTTGCTGCTGTACTCCTTCTCCGATG 143

      34 lnCysAsnValProGluTrpLeuProPhe.................Ala 44
           |||:::   |||:::      ||||||                 :::
     144 CCTGTGAGGAGCCACAA......CCATTTGAAGCTATGGAGCTCATTGGT 187

      45 ArgProThrAsnLeuThrAspAspPheGluPheProIleGlyThrTyrLe 61
         :::|||                     :::   ||||||      ::
     188 AAACCAAAACCCTAC..............TATGAGATTGGTGAACGAGT 222

      61 uAsnTyrGluCysArgProGlyTyrSerGlyArgPro.........PheS 75
         ::::|||:::|||:::   ||||||         |||
     223 AGATTATAAGTGTAAAAAAGGATACTTCTATATACCTCCTCTTGCCACCC 272

      75 erIleIleCysLeuLysAsnSerValTrpThrSer...AlaLysAspLys 90
           ||||||   :::|||   :::|||         :::   |||
     273 ATACTATTTGTGATCGGAATCATACATGGCTACCTGTCTCAGATGACGCC 322

      91 CysLysArgLysSerCysArgAsnProProAspProValAsnGlyMetAl 107
         |||   |||:::::|||           ||||||:::||||||:::||
     323 TGTTATAGAGAAACATGTCCATATATACGGGATCCTTTAAATGGCCAAGC 372

     107 aHisValIleLys...AspIleGlnPheGlySerGlnIleLysTyrSerC 123
         |         :::         :::||||||   |||:::   :::   |
     373 AGTCCCTGCAAATGGGACTTACGAGTTTGGTTATCAGATGCACTTTATTT 422

     123 ysProLysGlyTyrArgLeuIleGlySerSerSerAlaThrCysIleIle 139
         ||   :::||||||   ||||||||:::::         |||   :::
     423 GTAATGAGGGTTATTACTTAATTGGTGAAGAAATTCTATATTGTGAACTT 472

     140 SerGlyAsnThrValIleTrpAspAsnLysThrProValCysAspSerGl 156
         :::|||:::::::::||||||::::::|||   |||:::|||::::::
     473 AAAGGATCAGTAGCAATTTGGAGCGGTAAGCCCCCAATATGTGAAAAGGT 522

     156 uLeu 157
          |||
     523 TTTG 526

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-310-416A-11

seq_documentation_block:
; Sequence 11, Application US/08310416A
; Patent No. 5679546
;  GENERAL INFORMATION:
;    APPLICANT:  Jone-Long Ko et al.
;    TITLE OF INVENTION:  CHIMERIC PROTEINS WHICH BLOCK
;    TITLE OF INVENTION:  COMPLEMENT ACTIVATION
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/310,416A
;      FILING DATE:  22-SEP-1994
;      CLASSIFICATION:  435
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul T. Clark
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  06180/005001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  903 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-310-416A-11


alignment_scores:
           Quality:  289.00              Length:    173
             Ratio:  2.890                 Gaps:      6
 Percent Similarity:  57.803   Percent Identity: 39.884

alignment_block:
 US-10-031-904-8 x US-08-310-416A-11   ..

 Align seg 1/1  to: US-08-310-416A-11  from: 1  to: 903

       1 MetAlaProProValArgLeuGluArgProPheProSerArgArgPhePr 17
         |||   ||||||   |||   |||   ||||||||||||   ||||||||
       4 ATGGAGCCTCCCGGCCGCCGCGAGTGTCCCTTTCCTTCCTGGCGCTTTCC 53

      17 oGlyLeuLeuLeuAlaAlaLeuValLeuLeuLeuSerSerPheSerAspG 34
         ||||||||||||||||||:::||||||||||||   ||||||||||||
      54 TGGGTTGCTTCTGGCGGCCATGGTGTTGCTGCTGTACTCCTTCTCCGATG 103

      34 lnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThr 50
           |||:::   |||                |||||
     104 CCTGTGAGGAGCCA.....................CCAACA..... 123

      51 AspAspPheGlu..............................PhePr 56
            ||||||                               :::
     124 ......TTTGAAGCTATGGAGCTCATTGGTAAACCAAAACCCTACTATGA 167

      56 oIleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgP 73
          ||||||      ::::::|||:::|||:::   ||||||         |
     168 GATTGGTGAACGAGTAGATTATAAGTGTAAAAAAGGATACTTCTATATAC 217

      73 ro.........PheSerIleIleCysLeuLysAsnSerValTrpThrSer 86
         ||                  ||||||   :::|||   :::|||
     218 CTCCTCTTGCCACCCATACTATTTGTGATCGGAATCATACATGGCTACCT 267

      87 ...AlaLysAspLysCysLysArgLysSerCysArgAsnProProAspPr 102
            :::   |||   |||   |||:::::::|||              |||||
     268 GTCTCAGATGACGCCTGTTATAGAGAAACATGTCCATATATACGGGATCC 317

     102 oValAsnGlyMetAlaHisValIleLys...AspIleGlnPheGlySerG 118
         |:::||||||:::|||         :::          :::|||||||   |
     318 TTTAAATGGCCAAGCAGTCCCTGCAAATGGGACTTACGAGTTTGGTTATC 367

     118 lnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSerSer 134
         ||:::   :::   |||   :::|||||||   |||||||||::::::
     368 AGATGCACTTTATTTGTAATGAGGGTTATTACTTAATTGGTGAAGAAATT 417

     135 AlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysThrPr 151
              |||   ::::::|||:::::::::|||||:::::::|||   ||
     418 CTATATTGTGAACTTAAAGGATCAGTAGCAATTTGGAGCGGTAAGCCCCC 467

     151 oValCysAspSerGluLeu 157
         |:::|||::::::   |||
     468 AATATGTGAAAAGGTTTTG 486

 seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-888-171-11

seq_documentation_block:
; Sequence 11, Application US/08888171
; Patent No. 5851528
;  GENERAL INFORMATION:
;    APPLICANT:  Jone-Long, Ko
;    APPLICANT:  Higgins, Paul J.
;    APPLICANT:  Yeh, C. Grace
;    TITLE OF INVENTION:  METHODS OF INHIBITING COMPLEMENT
;    TITLE OF INVENTION:  ACTIVATION
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson, P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/888,171
;      FILING DATE:  03-JUL-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/310,416
;      FILING DATE:  22-SEP-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Freeman, John W.
;      REGISTRATION NUMBER:  29,066
;      REFERENCE/DOCKET NUMBER:  06180/005002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-507
;      TELEFAX:  617/542-890
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  903 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-888-171-11


alignment_scores:
           Quality:  289.00              Length:    173
             Ratio:  2.890                 Gaps:      6
 Percent Similarity:  57.803   Percent Identity: 39.884

alignment_block:
 US-10-031-904-8 x US-08-888-171-11   ..

 Align seg 1/1  to: US-08-888-171-11  from: 1  to: 903

       1 MetAlaProProValArgLeuGluArgProPheProSerArgArgPhePr 17
         |||   ||||||   |||   |||   ||||||||||||   ||||||||
       4 ATGGAGCCTCCCGGCCGCCGCGAGTGTCCCTTTCCTTCCTGGCGCTTTCC 53

      17 oGlyLeuLeuLeuAlaAlaLeuValLeuLeuLeuSerSerPheSerAspG 34
         ||||||||||||||||||:::||||||||||||   ||||||||||||
      54 TGGGTTGCTTCTGGCGGCCATGGTGTTGCTGCTGTACTCCTTCTCCGATG 103

      34 lnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThr 50
           |||:::   |||                |||||
     104 CCTGTGAGGAGCCA.....................CCAACA..... 123

      51 AspAspPheGlu..............................PhePr 56
            ||||||                               :::
     124 ......TTTGAAGCTATGGAGCTCATTGGTAAACCAAAACCCTACTATGA 167
```

```
      56 oIleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgP 73
         ||||||      ::::::|||:::|||:::   ||||||         |
     168 GATTGGTGAACGAGTAGATTATAAGTGTAAAAAAGGATACTTCTATATAC 217

      73 ro.........PheSerIleIleCysLeuLysAsnSerValTrpThrSer 86
         ||                 ||||||   :::|||   :::|||
     218 CTCCTCTTGCCACCCATACTATTTGTGATCGGAATCATACATGGCTACCT 267

      87 ...AlaLysAspLysCysLysArgLysSerCysArgAsnProProAspPr 102
            ::: . |||   |||   |||:::::::|||            |||||
     268 GTCTCAGATGACGCCTGTTATAGAGAAACATGTCCATATATACGGGATCC 317

     102 oValAsnGlyMetAlaHisValIleLys...AspIleGlnPheGlySerG 118
         |:::|||||||:::|||        :::       :::||||||   |
     318 TTTAAATGGCCAAGCAGTCCCTGCAAATGGGACTTACGAGTTTGGTTATC 367

     118 lnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSerSer 134
         ||:::   :::   |||   :::|||||||   |||||||||:::::
     368 AGATGCACTTTATTTGTAATGAGGGTTATTACTTAATTGGTGAAGAAATT 417

     135 AlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysThrPr 151
               |||   :::::::|||:::::::::|||||:::::::|||   ||
     418 CTATATTGTGAACTTAAAGGATCAGTAGCAATTTGGAGCGGTAAGCCCCC 467

     151 oValCysAspSerGluLeu 157
         |:::|||:::::   |||
     468 AATATGTGAAAAGGTTTTG 486

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-793-418-1

seq_documentation_block:
; Sequence 1, Application US/08793418C
; Patent No. 6130062
; GENERAL INFORMATION:
;  APPLICANT: The Austin Research Institute
;  TITLE OF INVENTION: Improvements in Production of Proteins in Host Cells
;  FILE REFERENCE: CALA-100
;  CURRENT APPLICATION NUMBER: US/08/793,418C
;  CURRENT FILING DATE: 1997-02-25
;  EARLIER APPLICATION NUMBER: PCT/AU95/00553
;  EARLIER FILING DATE: 1994-08-30
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1134
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  CD46  cDNA
;   OTHER INFORMATION: sequence with wild type SCR
US-08-793-418-1


alignment_scores:
            Quality:  289.00            Length:    173
              Ratio:  2.890               Gaps:      6
 Percent Similarity:  57.803   Percent Identity: 39.884

alignment_block:
 US-10-031-904-8 x US-08-793-418-1   ..

 Align seg 1/1  to: US-08-793-418-1  from: 1  to: 1134

       1 MetAlaProProValArgLeuGluArgProPheProSerArgArgPhePr 17
         |||   ||||||   |||   |||   ||||||||||||   ||||||||
       1 ATGGAGCCTCCCGGCCGCCGCGAGTGTCCCTTTCCTTCCTGGCGCTTTCC 50

      17 oGlyLeuLeuLeuAlaAlaLeuValLeuLeuLeuSerSerPheSerAspG 34
         ||||||||||||||||||:::||||||||||||   ||||||||||||
      51 TGGGTTGCTTCTGGCGGCCATGGTGTTGCTGCTGTACTCCTTCTCCGATG 100

      34 lnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThr 50
            |||:::   |||                  ||||||

     101 CCTGTGAGGAGCCA.....................CCAACA......... 120

      51 AspAspPheGlu.................................PhePr 56
               ||||||                                 :::
     121 ......TTTGAAGCTATGGAGCTCATTGGTAAACCAAAACCCTACTATGA 164

      56 oIleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgP 73
         ||||||      ::::::|||:::|||:::   ||||||         |
     165 GATTGGTGAACGAGTAGATTATAAGTGTAAAAAAGGATACTTCTATATAC 214

      73 ro.........PheSerIleIleCysLeuLysAsnSerValTrpThrSer 86
         ||                 ||||||   :::|||   :::|||
     215 CTCCTCTTGCCACCCATACTATTTGTGATCGGAATCATACATGGCTACCT 264

      87 ...AlaLysAspLysCysLysArgLysSerCysArgAsnProProAspPr 102
            ::: . |||   |||   |||:::::::|||            |||||
     265 GTCTCAGATGACGCCTGTTATAGAGAAACATGTCCATATATACGGGATCC 314

     102 oValAsnGlyMetAlaHisValIleLys...AspIleGlnPheGlySerG 118
         |:::|||||||:::|||        :::       :::||||||   |
     315 TTTAAATGGCCAAGCAGTCCCTGCAAATGGGACTTACGAGTTTGGTTATC 364

     118 lnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSerSer 134
         ||:::   :::   |||   :::|||||||   |||||||||:::::
     365 AGATGCACTTTATTTGTAATGAGGGTTATTACTTAATTGGTGAAGAAATT 414

     135 AlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysThrPr 151
               |||   :::::::|||:::::::::|||||:::::::|||   ||
     415 CTATATTGTGAACTTAAAGGATCAGTAGCAATTTGGAGCGGTAAGCCCCC 464

     151 oValCysAspSerGluLeu 157
         |:::|||:::::   |||
     465 AATATGTGAAAAGGTTTTG 483

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-793-418-3

seq_documentation_block:
; Sequence 3, Application US/08793418C
; Patent No. 6130062
; GENERAL INFORMATION:
;  APPLICANT: The Austin Research Institute
;  TITLE OF INVENTION: Improvements in Production of Proteins in Host Cells
;  FILE REFERENCE: CALA-100
;  CURRENT APPLICATION NUMBER: US/08/793,418C
;  CURRENT FILING DATE: 1997-02-25
;  EARLIER APPLICATION NUMBER: PCT/AU95/00553
;  EARLIER FILING DATE: 1994-08-30
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1134
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  CD46 cDNA
;   OTHER INFORMATION: subSCR3 variant
US-08-793-418-3


alignment_scores:
            Quality:  289.00            Length:    173
              Ratio:  2.890               Gaps:      6
 Percent Similarity:  57.803   Percent Identity: 39.884

alignment_block:
 US-10-031-904-8 x US-08-793-418-3   ..

 Align seg 1/1  to: US-08-793-418-3  from: 1  to: 1134

       1 MetAlaProProValArgLeuGluArgProPheProSerArgArgPhePr 17
         |||   ||||||   |||   |||   ||||||||||||   ||||||||
       1 ATGGAGCCTCCCGGCCGCCGCGAGTGTCCCTTTCCTTCCTGGCGCTTTCC 50
```

```
      17 oGlyLeuLeuLeuAlaAlaLeuValLeuLeuLeuSerSerPheSerAspG 34
         ||||||||||||||||||||:::|||||||||||||   ||||||||||||
      51 TGGGTTGCTTCTGGCGGCCATGGTGTTGCTGCTGTACTCCTTCTCCGATG 100

      34 lnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThr 50
           |||:::   |||                      ||||||
     101 CCTGTGAGGAGCCA.....................CCAACA......... 120

      51 AspAspPheGlu..............................PhePr 56
                 ||||||                           :::
     121 ......TTTGAAGCTATGGAGCTCATTGGTAAACCAAAACCCTACTATGA 164

      56 oIleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgP 73
          ||||||      :::::::|||:::|||:::   ||||||         |
     165 GATTGGTGAACGAGTAGATTATAAGTGTAAAAAAGGATACTTCTATATAC 214

      73 ro.........PheSerIleIleCysLeuLysAsnSerValTrpThrSer 86
         ||                   ||||||   :::|||   :::|||
     215 CTCCTCTTGCCACCCATACTATTTGTGATCGGAATCATACATGGCTACCT 264

      87 ...AlaLysAspLysCysLysArgLysSerCysArgAsnProProAspPr 102
            :::   |||   |||   |||::::::|||            |||||
     265 GTCTCAGATGACGCCTGTTATAGAGAAACATGTCCATATATACGGGATCC 314

     102 oValAsnGlyMetAlaHisValIleLys...AspIleGlnPheGlySerG 118
         |:::||||||:::|||         :::         :::||||||   |
     315 TTTAAATGGCCAAGCAGTCCCTGCAAATGGGACTTACGAGTTTGGTTATC 364

     118 lnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSerSer 134
         ||:::   :::   |||   :::||||||   ||||||||||:::::
     365 AGATGCACTTTATTTGTAATGAGGGTTATTACTTAATTGGTGAAGAAATT 414

     135 AlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysThrPr 151
               |||   ::::::|||:::::::::||||||:::::::|||   ||
     415 CTATATTGTGAACTTAAAGGATCAGTAGCAATTTGGAGCGGTAAGCCCCC 464

     151 oValCysAspSerGluLeu 157
         |:::|||::::::   |||
     465 AATATGTGAAAAGGTTTTG 483

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-528-057-41

seq_documentation_block:
; Sequence 41, Application US/08528057
; Patent No. 5846715
;  GENERAL INFORMATION:
;    APPLICANT:  PURCELL, Damian F. J.
;    APPLICANT:  RUSSELL, Sarah M.
;    APPLICANT:  McKENZIE, Ian F. C.
;    TITLE OF INVENTION:  CD46 VARIANTS
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/528,057
;      FILING DATE:  CONCURRENTLY HEREWITH
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/961,686
;      FILING DATE:  11-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/AU91/00199
;      FILING DATE:  10-MAY-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PK0133/90
;      FILING DATE:  11-MAY-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  17227/112 DACO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  41:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1247 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  83..1192
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  185..1192
US-08-528-057-41


alignment_scores:
            Quality:  289.00            Length:    173
              Ratio:  2.890               Gaps:      6
 Percent Similarity:  57.803   Percent Identity: 39.884

alignment_block:
 US-10-031-904-8 x US-08-528-057-41   ..

 Align seg 1/1  to: US-08-528-057-41  from: 1  to: 1247

       1 MetAlaProProValArgLeuGluArgProPheProSerArgArgPhePr 17
         |||   ||||||   |||   |||   ||||||||||||   ||||||||
      83 ATGGAGCCTCCCGGCCGCCGCGAGTGTCCCTTTCCTTCCTGGCGCTTTCC 132

      17 oGlyLeuLeuLeuAlaAlaLeuValLeuLeuLeuSerSerPheSerAspG 34
         ||||||||||||||||||||:::|||||||||||||   ||||||||||||
     133 TGGGTTGCTTCTGGCGGCCATGGTGTTGCTGCTGTACTCCTTCTCCGATG 182

      34 lnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThr 50
           |||:::   |||                      ||||||
     183 CCTGTGAGGAGCCA.....................CCAACA......... 202

      51 AspAspPheGlu..............................PhePr 56
                 ||||||                           :::
     203 ......TTTGAAGCTATGGAGCTCATTGGTAAACCAAAACCCTACTATGA 246

      56 oIleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgP 73
          ||||||      :::::::|||:::|||:::   ||||||         |
     247 GATTGGTGAACGAGTAGATTATAAGTGTAAAAAAGGATACTTCTATATAC 296

      73 ro.........PheSerIleIleCysLeuLysAsnSerValTrpThrSer 86
         ||                   ||||||   :::|||   :::|||
     297 CTCCTCTTGCCACCCATACTATTTGTGATCGGAATCATACATGGCTACCT 346

      87 ...AlaLysAspLysCysLysArgLysSerCysArgAsnProProAspPr 102
            :::   |||   |||   |||::::::|||            |||||
     347 GTCTCAGATGACGCCTGTTATAGAGAAACATGTCCATATATACGGGATCC 396

     102 oValAsnGlyMetAlaHisValIleLys...AspIleGlnPheGlySerG 118
         |:::||||||:::|||         :::         :::||||||   |
     397 TTTAAATGGCCAAGCAGTCCCTGCAAATGGGACTTACGAGTTTGGTTATC 446

     118 lnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSerSer 134
         ||:::   :::   |||   :::||||||   ||||||||||:::::
     447 AGATGCACTTTATTTGTAATGAGGGTTATTACTTAATTGGTGAAGAAATT 496

     135 AlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysThrPr 151
```

```
              |||   ::::::|||:::::::::|||||::::::|||   ||
      497 CTATATTGTGAACTTAAAGGATCAGTAGCAATTTGGAGCGGTAAGCCCCC 546

      151 oValCysAspSerGluLeu 157
          |:::|||::::::   |||
      547 AATATGTGAAAAGGTTTTG 565

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-528-057-45

seq_documentation_block:
; Sequence 45, Application US/08528057
; Patent No. 5846715
;  GENERAL INFORMATION:
;    APPLICANT:  PURCELL, Damian F. J.
;    APPLICANT:  RUSSELL, Sarah M.
;    APPLICANT:  McKENZIE, Ian F. C.
;    TITLE OF INVENTION:  CD46 VARIANTS
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/528,057
;      FILING DATE:  CONCURRENTLY HEREWITH
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/961,686
;      FILING DATE:  11-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/AU91/00199
;      FILING DATE:  10-MAY-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PK0133/90
;      FILING DATE:  11-MAY-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  17227/112 DACO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  45:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1304 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  94..1065
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  196..1065
US-08-528-057-45


alignment_scores:
            Quality:  289.00              Length:    173
              Ratio:  2.890                Gaps:      6
 Percent Similarity:  57.803   Percent Identity: 39.884

alignment_block:
 US-10-031-904-8 x US-08-528-057-45    ..
```

```
 Align seg 1/1  to: US-08-528-057-45  from: 1  to: 1304

       1 MetAlaProProValArgLeuGluArgProPheProSerArgArgPhePr 17
         |||   ||||||   |||   |||   ||||||||||||   ||||||||
      94 ATGGAGCCTCCCGGCCGCCGCGAGTGTCCCTTTCCTTCCTGGCGCTTTCC 143

      17 oGlyLeuLeuLeuAlaAlaLeuValLeuLeuLeuSerSerPheSerAspG 34
         ||||||||||||||||||:::|||||||||||   ||||||||||||
     144 TGGGTTGCTTCTGGCGGCCATGGTGTTGCTGCTGTACTCCTTCTCCGATG 193

      34 lnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThr 50
         |||:::   |||                    ||||||
     194 CCTGTGAGGAGCCA.....................CCAACA......... 213

      51 AspAspPheGlu.............................PhePr 56
               ||||||                             :::
     214 ......TTTGAAGCTATGGAGCTCATTGGTAAACCAAAACCCTACTATGA 257

      56 oIleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgP 73
          ||||||      ::::::|||:::|||:::   ||||||         |
     258 GATTGGTGAACGAGTAGATTATAAGTGTAAAAAAGGATACTTCTATATAC 307

      73 ro.........PheSerIleIleCysLeuLysAsnSerValTrpThrSer 86
         ||                  ||||||   :::|||   :::|||
     308 CTCCTCTTGCCACCCATACTATTTGTGATCGGAATCATACATGGCTACCT 357

      87 ...AlaLysAspLysCysLysArgLysSerCysArgAsnProProAspPr 102
            :::  |||   |||   |||:::::::|||              |||||
     358 GTCTCAGATGACGCCTGTTATAGAGAAACATGTCCATATATACGGGATCC 407

     102 oValAsnGlyMetAlaHisValIleLys...AspIleGlnPheGlySerG 118
         |:::||||||:::|||         :::         :::||||||   |
     408 TTTAAATGGCCAAGCAGTCCCTGCAAATGGGACTTACGAGTTTGGTTATC 457

     118 lnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSerSer 134
         ||:::   :::   |||   :::|||||||   |||||||||:::::
     458 AGATGCACTTTATTTGTAATGAGGGTTATTACTTAATTGGTGAAGAAATT 507

     135 AlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysThrPr 151
               |||   :::::::|||:::::::::||||||:::::::|||   ||
     508 CTATATTGTGAACTTAAAGGATCAGTAGCAATTTGGAGCGGTAAGCCCCC 557

     151 oValCysAspSerGluLeu 157
         |:::|||::::::   |||
     558 AATATGTGAAAAGGTTTTG 576

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-458-084-3

seq_documentation_block:
; Sequence 3, Application US/08458084
; Patent No. 5624837
;  GENERAL INFORMATION:
;    APPLICANT:  Fodor, William L
;    APPLICANT:  Rollins, Scott
;    APPLICANT:  Squinto, Stephen P
;    TITLE OF INVENTION:  Chimeric Complement
;    TITLE OF INVENTION:  Inhibitor Proteins
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maurice M. Klee
;      STREET:  1951 Burr Street
;      CITY:  Fairfield
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06430
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 750 Kb storage
;      COMPUTER:  Dell 486/50
;      OPERATING SYSTEM:  DOS 6.2
;      SOFTWARE:  WordPerfect 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,084
```

```
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Klee, Maurice M.
;      REGISTRATION NUMBER:  30,399
;      REFERENCE/DOCKET NUMBER:  ALX-120
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203) 255-1400
;      TELEFAX:  (203) 254-1101
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1530 bases
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Double
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  cDNA to mRNA
;      DESCRIPTION:  MCP (CD46) full length cDNA
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    PUBLICATION INFORMATION:
;      AUTHORS:  Lublin, D.M.
;      AUTHORS:  Liszewski,M.K.
;      AUTHORS:  Post, T.W.
;      AUTHORS:  Arce, M.A.
;      AUTHORS:  LeBeau, M.M.
;      AUTHORS:  Rebentisch, M.B.
;      AUTHORS:  Lemons, R.S.
;      AUTHORS:  Seya, T.
;      AUTHORS:  Atkinson, J.P.
;      TITLE:  Molecular cloning and Chromosomal
;      TITLE:  Localization of Membrane Cofactor
;      TITLE:  Protein (MCP): Evidence for Inclusion
;      TITLE:  in the Multi-Gene Family of
;      TITLE:  Complement-Regulatory Proteins.
;      JOURNAL:  Journal of Experimental Medicine
;      VOLUME:  168
;      PAGES:  181-194
;      DATE:  1988
US-08-458-084-3


alignment_scores:
           Quality:  289.00            Length:    173
             Ratio:  2.890               Gaps:      6
 Percent Similarity:  57.803   Percent Identity: 39.884

Alignment_block:
US-10-031-904-8 x US-08-458-084-3   ..

Align seq 1/1  to: US-08-458-084-3  from: 1  to: 1530

      1 MetAlaProProValArgLeuGluArgProPheProSerArgArgPhePr 17
        |||   ||||||   |||   |||   ||||||||||||   ||||||||
     44 ATGGAGCCTCCCGGCCGCCGCGAGTGTCCCTTTCCTTCCTGGCGCTTTCC 93

     17 oGlyLeuLeuLeuAlaAlaLeuValLeuLeuLeuSerSerPheSerAspG 34
        ||||||||||||||||||:::||||||||||||   ||||||||||||
     94 TGGGTTGCTTCTGGCGGCCATGGTGTTGCTGCTGTACTCCTTCTCCGATG 143

     34 lnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThr 50
          |||:::   |||                    ||||||
    144 CCTGTGAGGAGCCA.....................CCAACA..... 163

     51 AspAspPheGlu.................................PhePr 56
            ||||||                                  :::
    164 ......TTTGAAGCTATGGAGCTCATTGGTAAACCAAAACCCTACTATGA 207

     56 oIleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgP 73
        ||||||      :::::::|||:::|||:::   ||||||         |
    208 GATTGGTGAACGAGTAGATTATAAGTGTAAAAAAGGATACTTCTATATAC 257

     73 ro.........PheSerIleIleCysLeuLysAsnSerValTrpThrSer 86
        ||                    ||||||   :::|||   :::|||
    258 CTCCTCTTGCCACCCATACTATTTGTGATCGGAATCATACATGGCTACCT 307

     87 ...AlaLysAspLysCysLysArgLysSerCysArgAsnProProAspPr 102
           :::   |||   |||   |||:::::::|||                |||||
    308 GTCTCAGATGACGCCTGTTATAGAGAAACATGTCCATATATACGGGATCC 357

    102 oValAsnGlyMetAlaHisValIleLys...AspIleGlnPheGlySerG 118
        |:::||||||:::|||         :::         :::|||||   |
    358 TTTAAATGGCCAAGCAGTCCCTGCAAATGGGACTTACGAGTTTGGTTATC 407

    118 lnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSerSer 134
        ||:::   :::   |||   :::||||||   |||||||||:::::
    408 AGATGCACTTTATTTGTAATGAGGGTTATTACTTAATTGGTGAAGAAATT 457

    135 AlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysThrPr 151
            |||   ::::::|||:::::::::|||||||:::::|||   ||
    458 CTATATTGTGAACTTAAAGGATCAGTAGCAATTTGGAGCGGTAAGCCCCC 507

    151 oValCysAspSerGluLeu 157
        |:::|||::::::   |||
    508 AATATGTGAAAAGGTTTTG 526

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-205-508-3

seq_documentation_block:
; Sequence 3, Application US/08205508
; Patent No. 5627264
;  GENERAL INFORMATION:
;    APPLICANT:  Fodor, William L
;    APPLICANT:  Rollins, Scott
;    APPLICANT:  Squinto, Stephen P
;    TITLE OF INVENTION:  Chimeric Complement
;    TITLE OF INVENTION:  Inhibitor Proteins
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maurice M. Klee
;      STREET:  1951 Burr Street
;      CITY:  Fairfield
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06430
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 750 Kb storage
;      COMPUTER:  Dell 486/50
;      OPERATING SYSTEM:  DOS 6.2
;      SOFTWARE:  WordPerfect 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/205,508
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Klee, Maurice M.
;      REGISTRATION NUMBER:  30,399
;      REFERENCE/DOCKET NUMBER:  ALX-120
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203) 255-1400
;      TELEFAX:  (203) 254-1101
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1530 bases
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Double
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  cDNA to mRNA
;      DESCRIPTION:  MCP (CD46) full length cDNA
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    PUBLICATION INFORMATION:
;      AUTHORS:  Lublin, D.M.
```

```
;       AUTHORS:  Liszewski,M.K.
;       AUTHORS:  Post, T.W.
;       AUTHORS:  Arce, M.A.
;       AUTHORS:  LeBeau, M.M.
;       AUTHORS:  Rebentisch, M.B.
;       AUTHORS:  Lemons, R.S.
;       AUTHORS:  Seya, T.
;       AUTHORS:  Atkinson, J.P.
;       TITLE:  Molecular cloning and Chromosomal
;       TITLE:  Localization of Membrane Cofactor
;       TITLE:  Protein (MCP): Evidence for Inclusion
;       TITLE:  in the Multi-Gene Family of
;       TITLE:  Complement-Regulatory Proteins.
;       JOURNAL:  Journal of Experimental Medicine
;       VOLUME:  168
;       PAGES:  181-194
;       DATE:  1988
US-08-205-508-3


Alignment_scores:
            Quality:  289.00              Length:    173
              Ratio:  2.890                 Gaps:      6
 Percent Similarity:  57.803   Percent Identity: 39.884

alignment_block:
 US-10-031-904-8 x US-08-205-508-3   ..

 Align seg 1/1  to: US-08-205-508-3  from: 1  to: 1530

       1 MetAlaProProValArgLeuGluArgProPheProSerArgArgPhePr 17
         |||   ||||||   |||   |||   ||||||||||||   ||||||||
      44 ATGGAGCCTCCCGGCCGCCGCGAGTGTCCCTTTCCTTCCTGGCGCTTTCC 93

      17 oGlyLeuLeuLeuAlaAlaLeuValLeuLeuLeuSerSerPheSerAspG 34
         ||||||||||||||||||:::||||||||||||   ||||||||||||
      94 TGGGTTGCTTCTGGCGGCCATGGTGTTGCTGCTGTACTCCTTCTCCGATG 143

      34 lnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThr 50
           |||:::   |||                 ||||||
     144 CCTGTGAGGAGCCA.....................CCAACA......... 163

      51 AspAspPheGlu.................................PhePr 56
                 ||||||                              :::
     164 ......TTTGAAGCTATGGAGCTCATTGGTAAACCAAAACCCTACTATGA 207

      56 oIleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgP 73
         ||||||      ::::::|||:::|||:::   ||||||        |
     208 GATTGGTGAACGAGTAGATTATAAGTGTAAAAAAGGATACTTCTATATAC 257

      73 ro.........PheSerIleIleCysLeuLysAsnSerValTrpThrSer 86
         ||                  ||||||   :::|||   :::|||
     258 CTCCTCTTGCCACCCATACTATTTGTGATCGGAATCATACATGGCTACCT 307

      87 ...AlaLysAspLysCysLysArgLysSerCysArgAsnProProAspPr 102
            :::   |||   |||   |||:::::::|||             |||||
     308 GTCTCAGATGACGCCTGTTATAGAGAAACATGTCCATATATACGGGATCC 357

     102 oValAsnGlyMetAlaHisValIleLys...AspIleGlnPheGlySerG 118
         |:::|||||:::|||         :::         :::||||||   |
     358 TTTAAATGGCCAAGCAGTCCCTGCAAATGGGACTTACGAGTTTGGTTATC 407

     118 lnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSerSer 134
         ||:::   :::   |||   :::||||||   |||||||||:::::
     408 AGATGCACTTTATTTGTAATGAGGGTTATTACTTAATTGGTGAAGAAATT 457

     135 AlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysThrPr 151
               |||   :::::::|||:::::::::||||||:::::::|||   ||
     458 CTATATTGTGAACTTAAAGGATCAGTAGCAATTTGGAGCGGTAAGCCCCC 507

     151 oValCysAspSerGluLeu 157
         |:::|||::::::   |||
     508 AATATGTGAAAAGGTTTTG 526
```

```
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-482-148-8

seq_documentation_block:
; Sequence 8, Application US/08482148
; Patent No. 5847082
;  GENERAL INFORMATION:
;    APPLICANT:  Rother, Russell
;    APPLICANT:  Rollins, Scott
;    APPLICANT:  Squinto, Stephen P
;    TITLE OF INVENTION:  Terminal Complement
;    TITLE OF INVENTION:  Inhibitor Fusion Genes and Proteins
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seth A. Fidel
;      STREET:  Alexion Pharmaceuticals, 25 Science Park
;      CITY:  New Haven
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06511
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.4 Mb storage
;      COMPUTER:  IBM compatible (Pentium)
;      OPERATING SYSTEM:  Windows 95 under MS DOS
;      SOFTWARE:  Microsoft Word for Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,148
;      FILING DATE:  6/07/95
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fidel, Seth A.
;      REGISTRATION NUMBER:  38,449
;      REFERENCE/DOCKET NUMBER:  ALX-129.1 Div
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203) 776 1790
;      TELEFAX:  (203) 772 3655
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1530 base pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Double
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  cDNA to mRNA
;      DESCRIPTION:  MCP (CD46) full length cDNA
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    PUBLICATION INFORMATION:
;      AUTHORS:  Lublin, D.M.
;      AUTHORS:  Liszewski,M.K.
;      AUTHORS:  Post, T.W.
;      AUTHORS:  Arce, M.A.
;      AUTHORS:  LeBeau, M.M.
;      AUTHORS:  Rebentisch, M.B.
;      AUTHORS:  Lemons, R.S.
;      AUTHORS:  Seya, T.
;      AUTHORS:  Atkinson, J.P.
;      TITLE:  Molecular cloning and Chromosomal
;      TITLE:  Localization of Membrane Cofactor
;      TITLE:  Protein (MCP): Evidence for Inclusion
;      TITLE:  in the Multi-Gene Family of
;      TITLE:  Complement-Regulatory Proteins.
;      JOURNAL:  Journal of Experimental Medicine
;      VOLUME:  168
;      PAGES:  181-194
;      DATE:  1988
US-08-482-148-8


alignment_scores:
            Quality:  289.00              Length:    173
              Ratio:  2.890                 Gaps:      6
```

Percent Similarity: 57.803 Percent Identity: 39.884

alignment_block:
US-10-031-904-8 x US-08-482-148-8 ..

Align seg 1/1 to: US-08-482-148-8 from: 1 to: 1530

```
      1 MetAlaProProValArgLeuGluArgProPheProSerArgArgPhePr 17
        |||   ||||||   |||   |||   ||||||||||||   ||||||||
     44 ATGGAGCCTCCCGGCCGCCGCGAGTGTCCCTTTCCTTCCTGGCGCTTTCC 93

     17 oGlyLeuLeuLeuAlaAlaLeuValLeuLeuLeuSerSerPheSerAspG 34
        ||||||||||||||||||:::||||||||||||   ||||||||||||
     94 TGGGTTGCTTCTGGCGGCCATGGTGTTGCTGCTGTACTCCTTCTCCGATG 143

     34 lnCysAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThr 50
          |||:::   |||                    ||||||
    144 CCTGTGAGGAGCCA.....................CCAACA......... 163

     51 AspAspPheGlu.............................PhePr 56
             ||||||                             :::
    164 ......TTTGAAGCTATGGAGCTCATTGGTAAACCAAAACCCTACTATGA 207

     56 oIleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgP 73
         ||||||      ::::::|||:::|||:::   ||||||         |
    208 GATTGGTGAACGAGTAGATTATAAGTGTAAAAAAGGATACTTCTATATAC 257

     73 ro.......PheSerIleIleCysLeuLysAsnSerValTrpThrSer 86
        ||                  ||||||   :::|||   :::|||
    258 CTCCTCTTGCCACCCATACTATTTGTGATCGGAATCATACATGGCTACCT 307

     87 ...AlaLysAspLysCysLysArgLysSerCysArgAsnProProAspPr 102
           :::   |||   |||   |||:::::::|||           |||||
    308 GTCTCAGATGACGCCTGTTATAGAGAAACATGTCCATATATACGGGATCC 357

    102 oValAsnGlyMetAlaHisValIleLys...AspIleGlnPheGlySerG 118
        |:::||||||:::|||         :::         :::||||||   |
    358 TTTAAATGGCCAAGCAGTCCCTGCAAATGGGACTTACGAGTTTGGTTATC 407

    118 lnIleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSerSer 134
        ||:::   :::   |||   :::||||||   ||||||||||:::::
    408 AGATGCACTTTATTTGTAATGAGGGTTATTACTTAATTGGTGAAGAAATT 457

    135 AlaThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysThrPr 151
              |||   ::::::|||:::::::::||||||::::::|||   ||
    458 CTATATTGTGAACTTAAAGGATCAGTAGCAATTTGGAGCGGTAAGCCCCC 507

    151 oValCysAspSerGluLeu 157
        |:::|||::::::   |||
    508 AATATGTGAAAAGGTTTTG 526
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 9, 2002, 15:59:43 ; Search time 1838.24 Seconds
(without alignments)
7137.772 Million cell updates/sec

Title: US-10-031-904-30
Perfect score: 627
Sequence: 1 cggactcagaagggacttcc.....ataataaaaatcttaaccga 627

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

[redacted]tal number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	627	100.0	627	6	AX078362	AX078362 Sequence
2	352	56.1	6951	9	HSCR1	Y00816 Human mRNA
3	346.6	55.3	2376	9	HSCR1RS	X14362 Human CR1 m
4	346.6	55.3	2376	11	G28591	G28591 human STS S
5	343.4	54.8	1985	9	CHPCR1X	L24921 Pan troglod
6	343.4	54.8	6044	9	CHPCR1WT	L24920 Chimpanzee
7	328.8	52.4	1688	9	BABCORE	L77977 Papio cynoc
8	320	51.0	6000	9	BABCR1A	L39791 Papio hamad
9	316.8	50.5	945	9	BABCOREA	L77978 Papio cynoc
c 10	306.4	48.9	389	11	G25967	G25967 human STS E
c 11	306.4	48.9	389	11	G27827	G27827 human STS S
12	303.2	48.4	1731	9	CHPCR1Y	L24922 Pan troglod
13	267.4	42.6	591	6	A86593	A86593 Sequence 2
14	267.4	42.6	591	6	A86601	A86601 Sequence 10
15	267.4	42.6	591	6	A86603	A86603 Sequence 12
16	267.4	42.6	591	6	A86605	A86605 Sequence 14
17	267.4	42.6	591	6	A86607	A86607 Sequence 16
18	267.4	42.6	591	6	A86609	A86609 Sequence 18
19	267.4	42.6	605	6	AR029199	AR029199 Sequence
20	218.6	34.9	5420	6	AX333690	AX333690 Sequence
21	218.6	34.9	5420	9	HSCR1R	X05309 Human mRNA
22	216.6	34.5	309	9	HUMMCPLA	M73721 Human MCP-l
23	204.2	32.6	1680	10	D42115	D42115 Rattus norv
24	204.2	32.6	1724	10	RAT5I2	D42114 Rat 5I2 ant
25	204.2	32.6	1811	10	RATCRRY	L36532 Rat complem
26	202.2	32.2	1272	10	MUSCRRYPS	M34174 Mouse compl
27	201	32.1	1781	10	MUSCRYL	M23529 Mus musculu
28	199.6	31.8	1527	10	GPICRRP	M77760 Guinea pig
29	197	31.4	990	10	MUSXYG	M16179 Mouse X/Y p
30	194.8	31.1	145325	9	AL137789	AL137789 Human DNA
31	191.4	30.5	227	9	HUMCR1L2	M31231 Human compl
32	178.2	28.4	695	9	HUMCR1SF02	L17399 Human compl
33	178.2	28.4	328209	2	AL160003	AL160003 Homo sapi
34	177	28.2	227	9	HUMCR102	M31239 Human compl
35	174.6	27.8	2102	10	MUSCR2AB	M36470 Mouse compl
36	156.4	24.9	1304	6	AR063632	AR063632 Sequence
37	156.4	24.9	1339	9	HSMCP07	X59407 H.sapiens,
38	156.4	24.9	1429	9	HSMCP09	X59409 H.sapiens,
39	156.4	24.9	1432	9	HSMCP06	X59406 H.sapiens,
40	156.4	24.9	1477	6	A65266	A65266 Sequence 3
41	156.4	24.9	1522	9	HSMCP10	X59410 H.sapiens,
42	156.4	24.9	3225	9	HSMCP08	X59408 H.sapiens,
43	156.4	24.9	3318	9	HSMCP05	X59405 H.sapiens,
44	155.2	24.8	1500	9	S65879	S65879 MCP=complem
c 45	153.6	24.5	182742	2	AC084393	AC084393 Homo sapi

ALIGNMENTS

RESULT 1
AX078362
LOCUS AX078362 627 bp DNA linear PAT 22-FEB-2001
DEFINITION Sequence 30 from Patent WO0107612.
ACCESSION AX078362
VERSION AX078362.1 GI:13158031
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 627)
AUTHORS Au-Young,J., Bandman,O., Tang,Y.T., Yue,H., Azimzai,Y., Burford,N.,
Baughn,M.R., Lu,D.A., Hillman,J.L., Patterson,C. and Lal,P.
TITLE Receptors and associated proteins
JOURNAL Patent: WO 0107612-A 30 01-FEB-2001;
Incyte Genomics, Inc. (US)
FEATURES Location/Qualifiers
source 1. .627
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 103561CB1"
BASE COUNT 142 a 173 c 128 g 184 t

```
ORIGIN


  Query Match           100.0%;  Score 627;  DB 6;  Length 627;
  Best Local Similarity  100.0%;  Pred. No. 1e-173;
  Matches  627;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 cggactcagaagggacttccctgctcggctggctttcggtttctctgctcacctccggat 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CGGACTCAGAAGGGACTTCCCTGCTCGGCTGGCTTTCGGTTTCTCTGCTCACCTCCGGAT 60

Qy     61 aaatcacggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtccctttc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AAATCACGGGGTCTCCCGCGCCGCTCATGGCGCCTCCCGTCCGTCTCGAGCGTCCCTTTC 120

Qy    121 cttcccggcgctttcctgggttgcttctggcggccctggtgttgctgctgtcctccttct 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTTCCCGGCGCTTTCCTGGGTTGCTTCTGGCGGCCCTGGTGTTGCTGCTGTCCTCCTTCT 180

Qy    181 ccgatcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCGATCAATGCAATGTCCCGGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATG 240

Qy    241 actttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaa 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ACTTTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAA 300

Qy    301 gaccgttttctatcatctgcctaaaaaaactcagtctggacaagtgctaaggacaagtgca 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GACCGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACAAGTGCTAAGGACAAGTGCA 360

Qy    361 aacgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaag 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AACGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATGGCACATGTGATCAAAG 420

Qy    421 acatccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggtt 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ACATCCAGTTCGGATCCCAAATTAAATATTCTTGTCCTAAAGGATACCGACTCATTGGTT 480

Qy    481 cctcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctg 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CCTCGTCTGCCACATGCATCATCTCAGGCAACACTGTCATTTGGGATAATAAAACACCTG 540

Qy    541 tttgtgacagtgagttgaaatatgcattcctatttcttttaccgatacattctaatttttt 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TTTGTGACAGTGAGTTGAAATATGCATTCCTATTTCTTTTACCGATACATTCTAATTTTT 600

Qy    601 ctctggaataataaaaatcttaaccga 627
          |||||||||||||||||||||||||||
Db    601 CTCTGGAATAATAAAAATCTTAACCGA 627


RESULT   2
HSCR1
LOCUS       HSCR1                   6951 bp    mRNA    linear   PRI 23-MAR-1995
DEFINITION  Human mRNA for complement receptor type 1 (CR1, C3b/C4b receptor,
            CD35).
ACCESSION   Y00816
VERSION     Y00816.1  GI:30185
KEYWORDS    C3b/C4b complement component receptor; C3b/C4b receptor; CD35
            antigen; complement receptor; glycoprotein; membrane protein.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6951)
  AUTHORS   Klickstein,L.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-1988) Klickstein L.B., Room 617 Hunterian Bldg.,
            725 N. Wolfe St., Baltimore, MD 21205
REFERENCE   2  (bases 1 to 1531)
  AUTHORS   Klickstein,L.B., Bartow,T.J., Miletic,V., Rabson,L.D., Smith,J.A.
            and Fearon,D.T.
  TITLE     Identification of distinct C3b and C4b recognition sites in the
            human C3b/C4b receptor (CR1, CD35) by deletion mutagenesis
  JOURNAL   J. Exp. Med. 168 (5), 1699-1717 (1988)
  MEDLINE   89035992
COMMENT     This is the sequence of the F allotype of CR1.
            seq pos. 1532-6951 are already published see x05309 Data kindly
            reviewed (16/5/89) by Klickstein L.
FEATURES             Location/Qualifiers
     source          1. .6951
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /tissue_type="tonsil and HL-60"
                     /clone_lib="lambda S2T (ATCC #37546)."
     sig_peptide     28. .150
     CDS             28. .6147
                     /note="CR1 precursor protein"
                     /codon_start=1
                     /protein_id="CAA68755.1"
                     /db_xref="GI:30186"
                     /db_xref="SWISS-PROT:P17927"
                     /translation="MGASSPRSPEPVGPPAPGLPFCCGGSLLAVVVLLALPVAWGQCN
                     APEWLPFARPTNLTDEFEPPIGTYLNYECRPGYSGRPFSIICLKNSVWTGAKDRCRRK
                     SCRNPPDPVNGMVHVIKGIQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIWDNETPI
                     CDRIPCGLPPTITNGDFISTNRENFHYGSVVTYRCNPGSGGRKVFELVGEPSIYCTSN
                     DDQVGIWSGPAPQCIIPNKCTPPNVENGILVSDNRSLFSLNEVVEFRCQPGFVMKGPR
                     RVKCQALNKWEPELPSCSRVCQPPPDVLHAERTQRDKDNFSPGQEVFYSCEPGYDLRG
                     AASMRCTPQGDWSPAAPTCEVKSCDDFMGQLLNGRVLFPVNLQLGAKVDFVCDEGFQL
                     KGSSASYCVLAGMESLWNSSVPVCEQIFCPSPPVIPNGRHTGKPLEVFPFGKAVNYTC
                     DPHPDRGTSFDLIGESTIRCTSDPQGNGVWSSPAPRCGILGHCQAPDHFLFAKLKTQT
                     NASDFPIGTSLKYECRPEYYGRPFSITCLDNLVWSSPKDVCKRKSCKTPPDPVNGMVH
                     VITDIQVGSRINYSCTTGHRLIGHSSAECILSGNAAHWSTKPPICQRIPCGLPPTIAN
                     GDFISTNRENFHYGSVVTYRCNPGSGGRKVFELVGEPSIYCTSNDDQVGIWSGPAPQC
                     IIPNKCTPPNVENGILVSDNRSLFSLNEVVEFRCQPGFVMKGPRRVKCQALNKWEPEL
                     PSCSRVCQPPPDVLHAERTQRDKDNFSPGQEVFYSCEPGYDLRGAASMRCTPQGDWSP
                     AAPTCEVKSCDDFMGQLLNGRVLFPVNLQLGAKVDFVCDEGFQLKGSSASYCVLAGME
                     SLWNSSVPVCEQIFCPSPPVIPNGRHTGKPLEVFPFGKAVNYTCDPHPDRGTSFDLIG
                     ESTIRCTSDPQGNGVWSSPAPRCGILGHCQAPDHFLFAKLKTQTNASDFPIGTSLKYE
                     CRPEYYGRPFSITCLDNLVWSSPKDVCKRKSCKTPPDPVNGMVHVITDIQVGSRINYS
                     CTTGHRLIGHSSAECILSGNTAHWSTKPPICQRIPCGLPPTIANGDFISTNRENFHYG
                     SVVTYRCNLGSRGRKVFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNVENG
                     ILVSDNRSLFSLNEVVEFRCQPGFVMKGPRRVKCQALNKWEPELPSCSRVCQPPPEIL
                     HGEHTPSHQDNFSPGQEVFYSCEPGYDLRGAASLHCTPQGDWSPEAPRCAVKSCDDFL
                     GQLPHGRVLFPLNLQLGAKVSFVCDEGFRLKGSSVSHCVLVGMRSLWNNSVPVCEHIF
                     CPNPPAILNGRHTGTPSGDIPYGKEISYTCDPHPDRGMTFNLIGESTIRCTSDPHGNG
                     VWSSPAPRCELSVRAGHCKTPEQFPFASPTIPINDFEFPVGTSLNYECRPGYFGKMFS
                     ISCLENLVWSSVEDNCRRKSCGPPPEPFNGMVHINTDTQFGSTVNYSCNEGFRLIGSP
                     STTCLVSGNNVTWDKKAPICEIISCEPPPTISNGDFYSNNRTSFHNGTVVTYQCHTGP
                     DGEQLFELVGERSIYCTSKDDQVGVWSSPPPRCISTNKCTAPEVENAIRVPGNRSFFS
                     LTEIIRFRCQPGFVMVGSHTVQCQTNGRWGPKLPHCSRVCQPPPEILHGEHTLSHQDN
                     FSPGQEVFYSCEPSYDLRGAASLHCTPQGDWSPEAPRCTVKSCDDFLGQLPHGRVLLP
                     LNLQLGAKVSFVCDEGFRLKGRSASHCVLAGMKALWNSSVPVCEQIFCPNPPAILNGR
                     HTGTPFGDIPYGKEISYACDTHPDRGMTFNLIGESSIRCTSDPQGNGVWSSPAPRCEL
                     SVPAACPHPPKIQNGHYIGGHVSLYLPGMTISYTCDPGYLLVGKGFIFCTDQGIWSQL
                     DHYCKEVNCSFPLFMNGISKELEMKKVYHYGDYVTLKCEDGYTLEGSPWSQCQADDRW
                     DPPLAKCTSRAHDALIVGTLSGTIFFILLIIFLSWIILKHRKGNNAHENPKEVAIHLH
                     SQGGSSVHPRTLQTNEENSRVLP"
     mat_peptide     151. .6144
                     /product="mature CR1 protein (AA 1-1998)"
     misc_feature    151. .1500
                     /note="long homologous repeat A coding sequence"
     misc_feature    1501. .2850
                     /note="long homologous repeat B coding sequence"
     misc_feature    2851. .4209
                     /note="long homologous repeat C coding sequence"
     misc_feature    4210. .5565
                     /note="long homologous repeat D coding sequence"
BASE COUNT     1802 a   1680 c   1661 g   1808 t
ORIGIN


  Query Match           56.1%;  Score 352;  DB 9;  Length 6951;
  Best Local Similarity  84.0%;  Pred. No. 1.9e-92;
  Matches  410;  Conservative    0;  Mismatches   75;  Indels    3;  Gaps    1;
```

```
Qy      66 acggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtcccttccttcctcc 125
           | |||| |  |  | ||    |   ||   || ||| | |     |   |||  ||   ||
Db      28 ATGGGGGCCTCTTCTCCAAGAAGCCCGGAGCCTGTCGGGCCGCCGGCGCCCGGTCTCCCC 87

Qy     126 cggcgctttcctgggttgcttctggcggccctggtgttgct---gctgtcctccttctctcc 182
             |||     || |  || |||||||   ||||| ||||   |||| |     ||
Db      88 TTCTGCTGCGGAGGATCCCTGCTGGCGGTTGTGGTGCTGCTTGCGCTGCCGGTGGCCTGG 147

Qy     183 gatcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgac 242
           | |||||||||||| ||| ||||||||||||||||||||||||||||||||||||||||
Db     148 GGTCAATGCAATGCCCCAGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAG 207

Qy     243 tttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaaga 302
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208 TTTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGA 267

Qy     303 ccgttttctatcatctgcctaaaaaaactcagtctggacaagtgctaaggacaagtgcaaa 362
           ||||||||||||||||||||||||||||||||||||  |||||||||||| ||||| |
Db     268 CCGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACTGGTGCTAAGGACAGGTGCAGA 327

Qy     363 cgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagac 422
           |||||||||||||||||||||||||||||||||||||  |||||||||||| |
Db     328 CGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATGGTGCATGTGATCAAAGGC 387

Qy     423 atccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcc 482
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     388 ATCCAGTTCGGATCCCAAATTAAATATTCTTGTACTAAAGGATACCGACTCATTGGTTCC 447

Qy     483 tcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtt 542
           ||||||||||||||||||||||||  | ||||||||||||||||| |||||||| ||
Db     448 TCGTCTGCCACATGCATCATCTCAGGTGATACTGTCATTTGGGATAATGAAACACCTATT 507

Qy     543 tgtgacag 550
           ||||||||
Db     508 TGTGACAG 515
```

```
RESULT   3
HSCR1RS
LOCUS       HSCR1RS                 2376 bp    mRNA    linear   PRI 22-MAR-1995
DEFINITION  Human CR1 mRNA for C3b/C4b receptor secreted form.
ACCESSION   X14362 Y00812
VERSION     X14362.1  GI:30197
KEYWORDS    alternate splicing; C3b/C4b receptor; complement receptor;
            receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2376)
  AUTHORS   Hourcade,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-NOV-1988) Hourcade D., Howard Hughes Medical
            Institute, 660 S. Euclid St. Louis Mo, 63110, USA
REFERENCE   2  (bases 1 to 2376)
  AUTHORS   Hourcade,D., Miesner,D.R., Atkinson,J.P. and Holers,V.M.
  TITLE     Identification of an alternative polyadenylation site in the human
            C3b/C4b receptor (complement receptor type 1) transcriptional unit
            and prediction of a secreted form of complement receptor type 1
  JOURNAL   J. Exp. Med. 168 (4), 1255-1270 (1988)
  MEDLINE   89010527
COMMENT     The sequence overlaps with that reported by Klickstein et. al. in
            J. Exp. Med. 165:1095-1112(1987) x05309 and in J. Exp. Med.
            168:1699-1717(1988).
FEATURES             Location/Qualifiers
     source          1. .2376
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CR1-4"
                     /haplotype="CR1-A"
                     /cell_line="HL-60"
     CDS             <1. .1682
                     /note="CR1 receptor precursor (AA -16 to 543) (1 is 2nd
                     base in codon)"
                     /codon_start=3
                     /protein_id="CAA32541.1"
                     /db_xref="GI:736240"
                     /translation="SLLAVVVLLALPVAWGQCNAPEWLPFARPTNLTDEFEFPIGTYL
                     NYECRPGYSGRPFSIICLKNSVWTGAKDRCRRKSCRNPPDPVNGMVHVIKGIQFGSQI
                     KYSCTKGYRLIGSSSATCIISGDTVIWDNETPICDRIPCGLPPTITNGDFISTNRENF
                     HYGSVVTYRCNPGSGGRKVFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNV
                     ENGILVSDNRSLFSLNEVVEFRCQPGFVMKGPRRVKCQALNKWEPELPSCSRVCQPPP
                     DVLHAERTQRDKDNFSPGQEVFYSCEPGYDLRGAASMRCTPQGDWSPAAPTCEVKSCD
                     DFMGQLLNGRVLFPVNLQLGAKVDFVCDEGFQLKGSSASYCVLAGMESLWNSSVPVCE
                     QIFCPSPPVIPNGRHTGKPLEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIRCTSDPQ
                     GNGVWSSPAPRCGILGHCQAPDHFLFAKLKTQTNASDFPIGTSLKYECRPEYYGRPFS
                     ITCLDNLVWSSPKDVCKRKSCKTPPDPVNGMVHVITDIQVGSRINYSCTTG"
     sig_peptide     <1. .50
     mat_peptide     51. .1679
                     /product="mature CR1 receptor (AA 1-543); secreted form"
     polyA_site      2376
                     /note="polyA site"
BASE COUNT      633 a    549 c    568 g    626 t
ORIGIN
```

```
  Query Match             55.3%;  Score 346.6;  DB 9;  Length 2376;
  Best Local Similarity   90.5%;  Pred. No. 5.8e-91;
  Matches  370;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy     142 tgcttctggcggccctggtgttgctgctgtcctccttctccgatcaatgcaatgtcccgg 201
           ||||   ||  |   || || |   |||| |      ||  | |||||||||| ||| |
Db       7 TGCTGGCGGTTGTGGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAG 66

Qy     202 aatggcttccatttgccaggcctaccaacctaactgatgactttgagtttcccattggga 261
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      67 AATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCCCATTGGGA 126

Qy     262 catatctgaactatgaatgccgccctggttattccggaagaccgttttctatcatctgcc 321
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     127 CATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGACCGTTTTCTATCATCTGCC 186

Qy     322 taaaaaactcagtctggacaagtgctaaggacaagtgcaaacgtaaatcatgtcgtaatc 381
           ||||||||||||||||||  |||||||||||| ||||| ||||||||||||||||||||
Db     187 TAAAAAACTCAGTCTGGACTGGTGCTAAGGACAGGTGCAGACGTAAATCATGTCGTAATC 246

Qy     382 ctccagatcctgtgaatggcatggcacatgtgatcaaagacatccagttcggatcccaaa 441
           ||||||||||||||||||||||  |||||||||||||| ||||||||||||||||||||
Db     247 CTCCAGATCCTGTGAATGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAAA 306

Qy     442 ttaaatattcttgtcctaaaggataccgactcattggttcctcgtctgccacatgcatca 501
           ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db     307 TTAAATATTCTTGTACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCA 366

Qy     502 tctcaggcaacactgtcatttgggataataaaacacctgtttgtgacag 550
           |||||||  | |||||||||||||||||| ||||||| ||||||||||
Db     367 TCTCAGGTGATACTGTCATTTGGGATAATGAAACACCTATTTGTGACAG 415
```

```
RESULT   4
G28591
LOCUS       G28591                  2376 bp    DNA     linear   STS 11-JUL-1996
DEFINITION  human STS SHGC-35372, sequence tagged site.
ACCESSION   G28591
VERSION     G28591.1  GI:1408406
KEYWORDS    STS; STS sequence; primer; sequence tagged site.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2376)
  AUTHORS   Myers,R.M.
  JOURNAL   Unpublished
COMMENT
            Contact: Richard M. Myers
```

Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myers@shgc.stanford.edu

Primer A: TGAGTTGGCAGCAACATCTC
Primer B: CATACTCTTATATGTGCACTGCCC
STS size: 216
PCR Profile:
Initial incubation: 94 degrees C for 90 seconds

Denaturation: 94 degrees C for 15 seconds
Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600
Protocol:
Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/ul
Total Vol: 10 ul

Buffer:
MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from X14362
-- Washington University/Merck EST sequence.

```
FEATURES             Location/Qualifiers
     source          1. .2376
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="1"
     STS             1680. .1895
     primer_bind     1680. .1699
     primer_bind     complement(1872. .1895)
BASE COUNT      633 a    549 c    568 g    626 t
ORIGIN
```

Query Match 55.3%; Score 346.6; DB 11; Length 2376;
Best Local Similarity 90.5%; Pred. No. 5.8e-91;
Matches 370; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```
Qy    142 tgcttctggcggccctggtgttgctgctgtcctccttctccgatcaatgcaatgtcccgg 201
          ||||  ||  |  || || |  |||| |     ||  | ||||||||||| ||| |
Db      7 TGCTGGCGGTTGTGGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAG 66

Qy    202 aatggcttccatttgccaggcctaccaacctaactgatgactttgagtttcccattggga 261
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     67 AATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCCCATTGGGA 126

Qy    262 catatctgaactatgaatgccgccctggttattccggaagaccgttttctatcatctgcc 321
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    127 CATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGACCGTTTTCTATCATCTGCC 186

Qy    322 taaaaaactcagtctggacaagtgctaaggacaagtgcaaacgtaaatcatgtcgtaatc 381
          ||||||||||||||||  |||||||||||| ||||| ||||||||||||||||||||
Db    187 TAAAAAACTCAGTCTGGACTGGTGCTAAGGACAGGTGCAGACGTAAATCATGTCGTAATC 246

Qy    382 ctccagatcctgtgaatggcatggcacatgtgatcaaagacatccagttcggatcccaaa 441
          ||||||||||||||||||||||||  |||||||||||| ||||||||||||||||||||
Db    247 CTCCAGATCCTGTGAATGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAAA 306

Qy    442 ttaaatattcttgtcctaaaggataccgactcattggttcctcgtctgccacatgcatca 501
          ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    307 TTAAATATTCTTGTACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCA 366

Qy    502 tctcaggcaacactgtcatttgggataataaaacacctgtttgtgacag 550
          |||||||  | |||||||||||||||||| |||||||| ||||||||||
Db    367 TCTCAGGTGATACTGTCATTTGGGATAATGAAACACCTATTTGTGACAG 415
```

RESULT 5
CHPCR1X
LOCUS CHPCR1X 1985 bp mRNA linear PRI 30-OCT-1994
DEFINITION Pan troglodytes alternatively spliced CR1 (CR1) gene, partial cds.
ACCESSION L24921
VERSION L24921.1 GI:557726
KEYWORDS alternative splicing product; complement receptor 1.
SOURCE Pan troglodytes cDNA to mRNA.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 1985)
AUTHORS Birmingham,D.J., Shen,X.P., Hourcade,D., Nickells,M.W. and
Atkinson,J.P.
TITLE Primary sequence of an alternatively spliced form of CR1. Candidate
for the 75,000 M(r) complement receptor expressed on chimpanzee
erythrocytes
JOURNAL J. Immunol. 153 (2), 691-700 (1994)
MEDLINE 94292799

```
FEATURES             Location/Qualifiers
     source          1. .1985
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
                     /cell_line="EBV transformed"
                     /cell_type="B cell"
     sig_peptide     <1. .50
                     /gene="CR1"
     CDS             <1. .>1985
                     /gene="CR1"
                     /codon_start=3
                     /product="complement receptor 1"
                     /protein_id="AAA51439.1"
                     /db_xref="GI:557727"
                     /translation="SLLAVVVLLALPVAWGQCNAPEWLPFARPTNLTDEFEFPIGTYL
                     NYECRPGYYGRPFSIICLKNSVWTGAKDRCRRKSCRNPPDPVNGMVHVIKDIQFGSQI
                     KYSCTKGYRLIGSSSATCIISGDTVIWDNETPICDRIPCGLPPTITNGDFISTNRENF
                     HYGSVVTYRCNPGSGGRKVFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNV
                     ENGILVSDNRSLFSLNEVVEFRCQPGFVMKGPPRVKCQALNKWEPELPSCSRVCQPPP
                     DVLHAERTQRDKDNFSPGQEVFYSCEPGYDLRGAASLRCTPQGDWSPATPTCEVKSCD
                     DFMGQLLNGRVLFPVNLQLGAKVDFVCDEGFQLKGSSASYCVLAGMESLWNSSVPVCE
                     QIFCPNPPAILNGRHTGTPFGDIPYGKEISYACDTHPDRGMTFNLIGESSIRCTSDPQ
                     GNGVWSSPAPRCELSVPAACPHPPKIQNGHDIGGHVSLYLPGMTISYICDPGYLLVGK
                     GFIFCTDQGIWSQLDHYCKEVNCSFPLFMNGISKELEMKKVYHYGDYVTLKCEDGYTL
                     EGSPWSQCQADDRWDPPLAKCTSRAHDALIVGTSSGTIFFILLIIFLSWIILKHRKGN
                     NAHEKPKEVAIHLHSQGGSSVHPRTLQTNEENSRVLP"
     gene            1. .1985
                     /gene="CR1"
     mat_peptide     51. .1985
                     /gene="CR1"
                     /product="complement receptor 1"
BASE COUNT      521 a    476 c    478 g    510 t
ORIGIN
```

Query Match 54.8%; Score 343.4; DB 9; Length 1985;
Best Local Similarity 90.0%; Pred. No. 4.9e-90;
Matches 368; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
Qy    142 tgcttctggcggccctggtgttgctgctgtcctccttctccgatcaatgcaatgtcccgg 201
          ||||  ||  |  || || |  |||| |     ||  | ||||||||||| ||| |
Db      7 TGCTGGCGGTTGTGGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAG 66

Qy    202 aatggcttccatttgccaggcctaccaacctaactgatgactttgagtttcccattggga 261
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     67 AATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCCCATTGGGA 126

Qy    262 catatctgaactatgaatgccgccctggttattccggaagaccgttttctatcatctgcc 321
          ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
```

```
Db    127 CATATCTGAACTATGAATGCCGCCCTGGTTATTATGGAAGACCGTTTTCTATCATCTGCC 186

Qy    322 taaaaaactcagtctggacaagtgctaaggacaagtgcaaacgtaaatcatgtcgtaatc 381
          ||||||||||||||||||  |||||||||||| |||||| ||||||||||||||||||||
Db    187 TAAAAAACTCAGTCTGGACTGGTGCTAAGGACAGGTGCAGACGTAAATCATGTCGTAATC 246

Qy    382 ctccagatcctgtgaatggcatggcacatgtgatcaaagacatccagttcggatcccaaa 441
          ||||||||||||||||||||||||  |||||||||||||||||||| ||||||||||||
Db    247 CTCCAGATCCTGTGAATGGCATGGTGCATGTGATCAAAGACATCCAATTCGGATCCCAAA 306

Qy    442 ttaaatattcttgtcctaaaggataccgactcattggttcctcgtctgccacatgcatca 501
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    307 TTAAATATTCTTGTACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCA 366

Qy    502 tctcaggcaacactgtcatttgggataataaaacacctgtttgtgacag 550
          |||||||  | ||||||||||||||||||| |||||||| ||||||||||
Db    367 TCTCAGGTGATACTGTCATTTGGGATAATGAAACACCTATTTGTGACAG 415
```

RESULT 6
CHPCR1WT
LOCUS CHPCR1WT 6044 bp mRNA linear PRI 30-OCT-1994
DEFINITION Chimpanzee complement receptor type one (CR1) mRNA, partial cds.
ACCESSION L24920
VERSION L24920.1 GI:551564
KEYWORDS complement receptor 1.
SOURCE Pan troglodytes cDNA to mRNA.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 6044)
AUTHORS Birmingham,D.J., Shen,X.P., Hourcade,D., Nickells,M.W. and Atkinson,J.P.
TITLE Primary sequence of an alternatively spliced form of CR1. Candidate for the 75,000 M(r) complement receptor expressed on chimpanzee erythrocytes
JOURNAL J. Immunol. 153 (2), 691-700 (1994)
MEDLINE 94292799
FEATURES Location/Qualifiers
source 1. .6044
/organism="Pan troglodytes"
/db_xref="taxon:9598"
gene 1. .6044
/gene="CR1"
CDS <1. .>6044
/gene="CR1"
/codon_start=3
/product="complement receptor 1"
/protein_id="AAA51438.1"
/db_xref="GI:557725"
/translation="SLLAVVVLLALPVAWGQCNAPEWLPFARPTNLTDEFEFPIGTYL
NYECRPGYYGRPFSIICLKNSVWTGAKDRCRRKSCRNPPDPVNGMVHVIKDIQFGSQI
KYSCTKGYRLIGSSSATCIISGDTVIWDNETPICDRIPCGLPPTITNGDFISTNRENF
HYGSVVTYRCNPGSGGRKVFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNV
ENGILVSDNRSLFSLNEVVEFRCQPGFVMKGPPRVKCQALNKWEPELPSCSRVCQPPP
DVLHAERTQRDKDNFSPGQEVFYSCEPGYDLRGAASLRCTPQGDWSPATPTCEVKSCD
DFMGQLLNGRVLFPVNLQLGAKVDFVCDEGFQLKGSSASYCVLAGMESLWNSSVPVCE
QIFCPSPPVIPNGRHTGKPLEVFPFGKAVNYTCDPHPDRGTTFDLIGESTIRCTSDPQ
GNGVWSSPAPRCGILGHCQAPDHFLFAKLKTQTNASDFPIGTSLKYKCRPEYYGRPFS
ITCLDNLVWSSPKDVCKRKSCKTPPDPVNGMVHVITDIQVGSRINYSCTTGHRLIGHS
SAECILSGNAAHWSTKPPICQRIPCGLPPTIANGDFISTNRENFHYGSVVTYRCNPGS
RGRKVFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNVENGILVSDNRSLFS
LNEVVEFRCQPGFAMKGPRRVKCQALNKWEPELPSCSRVCQPPPDVLHAERTQRDKDN
FSPGQEVFYSCEPGYDLRGAASLRCTPQGDWSPAAPTCEVKSCDGFMGQLLNGRVLFP
VNLQLGAKVDFVCDEGFQLKGSSASYCVLAGMESLWNSSVPVCEQIFCPSPPVIPNGR
HTGKPLEVFPFGKAVNYTCDPHPDRGTTFDLIGESTIRCTSDPQGNGVWSSPAPRCGI
LGHCQAPDHFLFAKLKTQTNASDFPIGTSLKYECRPEYYGRPFSITCLDNLVWSSPKD
VCRRKSCKTPPDPVNGMVHVITDIQVGSRINYSCTTGHRLIGHSSAECILSGNSAHWS
TKPPICQRIPCGLPPTIANGDFISTNRENFHYGSVVTYRCNLGSRGRKVFELVGEPSI
YCTSNDDQVGIWSGPAPQCIIPNKCTPPNVENGILVSDNRSLFSLNEVVEFRCQPGFV
MKGPHRVKCQALNKWEPELPSCSRVCQPPPEILHGEHTPSHQDNFSPGQEVFYSCEPG
YDLRGAASLHCTPQGDWSPEAPRCAVKSCDDFLGQLPHGRVLFPLNLQLGAKVSFVCD
EGFRLKGSSVSHCVLVGMRSLWNNSVPVCEQIFCPNPPAILNGRHTGTPFGDIPYGKE
ISYTCDPHPDRGMTFNLIGESTIRCTSDPHGNGVWSSPAPRCELPVHAGHCKTPEQFP
FASPTIPINDFEFPVGTSLNYECRPGYFGKMFSISCLENLVWSSVEDNCRRKSCGPPP
EPFNGMVHINTDTQFGSTVNYSCNEGFRLIGSPSTTCLVSGNNVTWDKKAPICEIISC
EPPPTISNGDFYSNNRASFHNGTVVTYQCHTGPDGEQLFELVGERSIYCTSKDDQVGV
WSSPPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEIVRFRCQPGFVMVGSHTVQCQT
NGRWGPKLPHCSRVCQPPPEILHGEHTPSHQDNFSPGQEVFYSCEPGYDLRGAASLHC
TPQGDWSPEAPRCTVKSCDDFLGQLPHGRVLFPLNLQLGAKVSFVCDEGFRLKGRSAS
HCVLAGMKALWNSSVPVCEQIFCPNPPAILNGRHTGTPFGDIPYGKEISYACDTHPDR
GMTFNLIGESSIRCTSDPQGNGVWSSPAPRCELSVPAACPHPPKIQNGHDIGGHVSLY
LPGMTISYICDPGYLLVGKGFIFCTDQGIWSQLDHYCKEVNCSFPLFMNGISKELEMK
KVYHYGDYVTLKCEDGYTLEGSPWSQCQADDRWDPPLAKCTSRAHDALIVGTSSGTIF
FILLIIFLSWIILKHRKGNNAHEKPKEVAIHLHSQGGSSVHPRTLQTNEENSRVLP"
BASE COUNT 1561 a 1496 c 1462 g 1525 t
ORIGIN

```
Query Match 54.8%; Score 343.4; DB 9; Length 6044;
Best Local Similarity 90.0%; Pred. No. 6.2e-90;
Matches 368; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy    142 tgcttctggcggccctggtgttgctgctgtcctccttctccgatcaatgcaatgtcccgg 201
          ||||  ||  |  || || |  |||| |     ||  | ||||||||||| ||| |
Db      7 TGCTGGCGGTTGTGGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAG 66

Qy    202 aatggcttccatttgccaggcctaccaacctaactgatgactttgagtttcccattggga 261
          ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     67 AATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCCCATTGGGA 126

Qy    262 catatctgaactatgaatgccgccctggttattccggaagaccgttttctatcatctgcc 321
          |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||
Db    127 CATATCTGAACTATGAATGCCGCCCTGGTTATTATGGAAGACCGTTTTCTATCATCTGCC 186

Qy    322 taaaaaactcagtctggacaagtgctaaggacaagtgcaaacgtaaatcatgtcgtaatc 381
          ||||||||||||||||||  |||||||||||| |||||| ||||||||||||||||||||
Db .  187 TAAAAAACTCAGTCTGGACTGGTGCTAAGGACAGGTGCAGACGTAAATCATGTCGTAATC 246

Qy    382 ctccagatcctgtgaatggcatggcacatgtgatcaaagacatccagttcggatcccaaa 441
          ||||||||||||||||||||||||  |||||||||||||||||||| ||||||||||||
Db    247 CTCCAGATCCTGTGAATGGCATGGTGCATGTGATCAAAGACATCCAATTCGGATCCCAAA 306

Qy    442 ttaaatattcttgtcctaaaggataccgactcattggttcctcgtctgccacatgcatca 501
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    307 TTAAATATTCTTGTACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCA 366

Qy    502 tctcaggcaacactgtcatttgggataataaaacacctgtttgtgacag 550
          |||||||  | ||||||||||||||||||| |||||||| ||||||||||
Db    367 TCTCAGGTGATACTGTCATTTGGGATAATGAAACACCTATTTGTGACAG 415
```

RESULT 7
BABCORE
LOCUS BABCORE 1688 bp mRNA linear PRI 07-MAY-1996
DEFINITION Papio cynocephalus complement receptor mRNA, partial cds.
ACCESSION L77977
VERSION L77977.1 GI:1301608
KEYWORDS complement C3b; complement receptor; glycophosphatidylinositol-linked protein.
SOURCE Papio cynocephalus cDNA to mRNA.
ORGANISM Papio cynocephalus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Papio.
REFERENCE 1 (bases 1 to 1688)
AUTHORS Birmingham,D.J., Logar,C.M., Shen,X.-P. and Chen,W.
TITLE The baboon erythrocyte complement receptor is a glycophosphatidyl inositol-linked protein encoded by a homologue of the human CR1-like genetic element
JOURNAL Unpublished (1996)
FEATURES Location/Qualifiers
source 1. .1688
/organism="Papio cynocephalus"
/db_xref="taxon:9556"
/cell_type="erythrocyte"

```
                     /tissue_type="bone marrow"
                     /dev_stage="adult"
     mRNA            <1. .>1688
     CDS             <1. .1571
                     /note="homologue of human CR1-like genetic element"
                     /codon_start=3
                     /product="complement receptor"
                     /protein_id="AAA99004.1"
                     /db_xref="GI:1301609"
                     /translation="VVLQRMGASSPRSPEPVGPPAPRLLFCCGGSLLAVVVLLALPVA
                     WGQCNAPEQLPFARPTNLTDASEFPVGTYLKYECLPGYHGKPFSIICLKNSVWTSAKD
                     KCTRKSCRNPKDPVNGMVHVIKDIQFGSQINYSCNKGYRLIGSSSATCIISGNTVIWD
                     NETPICEIIPCGLPPTIANGDFISTSREYFPYGSVVTYRCNLGSGRKKLFELVGEPSI
                     YCTSKDDQVGIWSGPAPQCIIPNKCMPPNVENGVLVSVNRSLFSLNEVVEFRCQPGFV
                     MKGPRHVQCQALNKWEPELPSCSRVCQPPPEILHGEHTPSHQDFSPGQEVFYSCEPGY
                     DLRGAASLHCTPQGDWNPEAPICTVKSCDDFLGQLPHGRVLFPLNLQLGAKVSFVCDE
                     GFRLKGRFASHCVLAGMKALWNSSVPVCEQIFCPNPPAILNGRHIGAPLGDIPYGKEV
                     SYICDPHPDRGMTVNLIGESTIRCTSDPQGNGVWSSPAPRCELSVPAGANDALIVGTL
                     SGTIFVILFIIFLS"
BASE COUNT      413 a    427 c    411 g    437 t
ORIGIN

  Query Match             52.4%;  Score 328.8;  DB 9;  Length 1688;
  Best Local Similarity   80.7%;  Pred. No. 9.5e-86;
  Matches  397;  Conservative    0;  Mismatches   92;  Indels    3;  Gaps    1;

Qy      61 aaatcacggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtcccttc 120
           | |  | |||| |  |  | ||     |   ||   || ||  | |    |   |||  ||
Db      13 AGAGAATGGGGGCCTCTTCTCCAAGAAGCCCGGAGCCGGTTGGGCCGCCGGCGCCCCGTC 72

Qy     121 cttcccggcgctttcctgggttgcttctggcggccctggtgttgct---gctgtcctcct 177
               |    |||     || |  ||  ||||||   ||||| ||||   |||| |
Db      73 TCCTCTTCTGCTGCGGAGGATCCCTGTTGGCGGTTGTGGTGCTGCTCGCGCTGCCGGTGG 132

Qy     178 tctccgatcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactg 237
           ||  | ||||||||||| |||||||  ||||||||||||||||||||||||||||||||
Db     133 CCTGGGGTCAATGCAATGCCCCGGAACAGCTTCCATTTGCCAGGCCTACCAACCTAACTG 192

Qy     238 atgactttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccg 297
           ||| | ||||||||| ||||||||||||||| |||||||||| ||||||||||   |
Db     193 ATGCATCTGAGTTTCCCGTTGGGACATATCTGAAGTATGAATGCCTCCCTGGTTATCATG 252

Qy     298 gaagaccgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagt 357
           ||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     253 GAAAACCATTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACAAGTGCTAAGGACAAGT 312

Qy     358 gcaaacgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatca 417
           ||| |||||||||||||||||||||  ||||||||||||||||||  ||||||||||||
       313 GCACACGTAAATCATGTCGTAATCCTAAAGATCCTGTGAATGGCATGGTGCATGTGATCA 372

Qy     418 aagacatccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattg 477
           |||||||||||||||||||||||||||| |||||||||  ||||||||||||||||||||
Db     373 AAGACATCCAGTTCGGATCCCAAATTAATTATTCTTGTAATAAAGGATACCGACTCATTG 432

Qy     478 gttcctcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacac 537
           |||||||||| |||||||||||||||||||||||||| |||||||||||||||||| ||||||
Db     433 GTTCCTCGTCCGCCACATGCATCATCTCAGGCAATACTGTCATTTGGGATAATGAAACAC 492

Qy     538 ctgtttgtgaca 549
           || ||||||| |
Db     493 CTATTTGTGAGA 504


RESULT   8
BABCR1A
LOCUS       BABCR1A                 6000 bp    mRNA    linear   PRI 11-FEB-1995
DEFINITION  Papio hamadryas complement component receptor type 1 (CR1) mRNA,
            complete cds.
ACCESSION   L39791
VERSION     L39791.1  GI:662828
KEYWORDS    C3b/C4b complement component receptor; complement component
            receptor CR1; complement receptor 1.
SOURCE      Papio hamadryas (clone SPC-CYT3NC) cDNA to mRNA.
  ORGANISM  Papio hamadryas
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Papio.
REFERENCE   1  (bases 1 to 6000)
  AUTHORS   Clemenza,L., Subramanian,B.V., Nickells,M.W., Hourcade,D.E. and
            Atkinson,J.P.
  TITLE     Primary sequence of the baboon 200 kDa C3b/C4b receptor (CR1)
  JOURNAL   Unpublished (1995)
FEATURES             Location/Qualifiers
     source          1. .6000
                     /organism="Papio hamadryas"
                     /db_xref="taxon:9557"
                     /clone="SPC-CYT3NC"
                     /cell_line="26CB-1"
                     /cell_type="lymphoblastoid"
     gene            1. .6000
                     /gene="CR1"
     CDS             268. .>6000
                     /gene="CR1"
                     /function="acts as receptor for complement component
                     C3b/C4b"
                     /note="putative"
                     /codon_start=1
                     /product="complement component receptor type 1"
                     /protein_id="AAA62170.1"
                     /db_xref="GI:662829"
                     /translation="MVHVIKDIQFGSQINYSCTEGHRLIGSSSATCIISGNTVIWDNE
                     TPICEKISCGLPPTIDNGDFFSANKEYFHYGSVVTYRCNLGSGGRKLFELVGEPSIYC
                     TSNEDQVGIWSGPAPQCIIPNKCTPPNVENGILVSVNRSLFSLNEVVEFRCQPGFVMK
                     GPRRVQCQALNKWEPELPSCSRVCQPPPDVLHGERTQRDKDIFQTGQEVFYICEPGYD
                     LRGAASLRCTPQGDWSPAAPRCEVKSCDDSLGQLPNGRVLFPRSLQLGAKVDFVCDEG
                     FQLKGSSASYCVLAGMESLWNSSVPVCEQIFCPSPPVIPNGRHTGKPLEVFPFGKAVT
                     YTCDPHPDRGMTFDLIGESTIRCTSDPQGNGVWSSPAPRCGILGHCKAPDHFLFAKLK
                     TQTNASDFPIGTSLKYECRPEYYGRPFSITCLDNLEWSSPKDVCKRKSCKTPPDPVNG
                     MVHVITDIQVGSRINYSCTTGHRLIGHSSAECVTSGNTAHWSTKPPICQRIPCGLPPP
                     IANGDFISTNREYFHYGSVVTYRCNLGSGRKKLFELVGEPSIYCTSKDDQVGIWSGPA
                     PQCIIPNKCMPPNVENGILVSVNRSLFSLNEVVEFRCQPGFVMKGPRRVQCQALNKWE
                     PELPSCSRVCQPPPDVLHGERTQRDKDIFQPGQEVFYICEPGYDLRGAASLRCTPQGD
                     WSPAAPRCEVKSCDDSLGQLPNGRVLFPRSLQLGAKVDFVCDEGFQLKGSSASYCVLA
                     GMESLWNSSVPVCEQIFCPSPPVIPNGRHTGKPLEVFPFGKAVTYTCDPHPDRGMTFD
                     LIGESTIRCTSDPQGNGVWSSRAPRCGILGHCKAPDHFLFAKLKTQTNASDFPIGTSL
                     KYECRPEYYGKPFSITCLDNLVWSSPKDVCKRKSCKTPPDPVNGMVHVITDIQVGSRI
                     NYSCTTGHRLIGHSSAECIISGNTAHWSTKPPICQRIPCGLPPPIANGDFISTNREYF
                     HYGSVVTYRCNLGSGRKKLFELVGEPSIYCTSKDDQVGIWSGPAPQCIIPNKCMPPNV
                     ENGILVSVNRSLFSLNEVVEFRCQPGFVMKGPRRVQCQALNKWEPELPSCSRVCQPPP
                     EILHGEHTPSHQDKFSPGQEVFYSCEPGYDLRGAASLHCTPQGDWSPEAPRCAVKSCD
                     DFLGQLHHGRVLVPFNLQLGAKVSFVCDEGFRLKGSSVSHCVLVGMRSLWNNSVPVCE
                     HIFCPNPPAILNGRHTGALLGDIPYGKEISYTCDPHRDRGMTFNLIGESTIRCTSDLQ
                     GNGVWSSPAPRCELSVRAGHCKTPEQFPFASPTIPINDFEFPVGTSLNYECHPGYFGR
                     MFSISCLENLVWSSVEDNCRRKSCGTPPEPFNGMVHINTDTQFGSTVNYSCNEGFRLI
                     GSPSTTCLVSGNNVTWDKEAPICEIISCKPPPTISNGDFYSNNRTSFHSGTVVTYQCH
                     TGPDGEQLFELVGERSIYCTSKDDQVGAWSSPPPRCISTNKCTAPEVKNGIRVPGNRS
                     FFSLNEIVRFRCQPGFVMVGSHTVQCQTNNRWGPKLPHCSRVCQPPPEILHGEHTPSH
                     QDKFSPGQEVFYSCEPGYDLRGAASLHCTPQGDWSPEAPICTVKSCDEFLGQLPHGRV
                     LSPLNLQLGAKVSFVCDEGFRLKGRSASHCVLAGMKALWNSSVPVCEQIFCPNPPAIL
                     NGRHTGTPLGDIPYGKEVSYTCDPHPDRGMTFNLIGESTIRCTSDLQGNGVWSSPAPR
                     CELSVPAACPHPPKIQNGHYIGGHVSLYLPGMTIGYICDPGYLLVGKGIIFCTDQGIW
                     SQLDHYCKEVNCSFPQFMNGISKELEMKKVYHYGDYVTLECEDGYALEGSPWSQCQAD
                     DRWDPPLAICTSRARDALIVGTLSGMIFVILFIIFLSWIILKYRKGNNAHEKPKEVAI
                     HIHSQGGSSVH"
BASE COUNT     1540 a   1461 c   1465 g   1534 t
ORIGIN

  Query Match             51.0%;  Score 320;  DB 9;  Length 6000;
  Best Local Similarity   86.5%;  Pred. No. 4.9e-83;
  Matches  353;  Conservative    0;  Mismatches   55;  Indels    0;  Gaps    0;

Qy     142 tgcttctggcggccctggtgttgctgctgtcctccttctccgatcaatgcaatgtcccgg 201
           ||||   ||  |   || || |  |||| |      ||  | ||||||||||| |||||
Db       8 TGCTGGCGGTTGTGGTGCTGCTCGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCGG 67
```

```
Qy    202 aatggcttccatttgccaggcctaccaacctaactgatgactttgagtttcccattggga 261
          ||  |||||||||||||||||||||  | |||| ||||||   ||||||| |||||||||
Db     68 AACAGCTTCCATTTGCCAGGCCTACTGAACTAATTGATGAGCCTGAGTTTTCCATTGGGA 127

Qy    262 catatctgaactatgaatgccgccctggttattccggaagaccgttttctatcatctgcc 321
          || ||||||| |||||||| ||||||||||||   |||||||| ||||||||||||||||
Db    128 CACATCTGAAGTATGAATGTCGCCCTGGTTATTATGGAAGACCATTTTCTATCATCTGCC 187

Qy    322 taaaaaactcagtctggacaagtgctaaggacaagtgcaaacgtaaatcatgtcgtaatc 381
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    188 TAAAAAACTCAGTCTGGACAAGTGCTAAGGACAAGTGCATACGTAAATCATGTCGTAATC 247

Qy    382 ctccagatcctgtgaatggcatggcacatgtgatcaaagacatccagttcggatcccaaa 441
          ||  ||||||||||||||||||||  |||||||||||||||||||||||||||||||| |
Db    248 CTAGAGATCCTGTGAATGGCATGGTGCATGTGATCAAAGACATCCAGTTCGGATCCCAGA 307

Qy    442 ttaaatattcttgtcctaaaggataccgactcattggttcctcgtctgccacatgcatca 501
          |||| ||||||||| || ||||| ||||||||||||||||||| ||||||||||||||||
          308 TTAATTATTCTTGTACTGAAGGACACCGACTCATTGGTTCCTCATCTGCCACATGCATCA 367

Qy    502 tctcaggcaacactgtcatttgggataataaaacacctgtttgtgaca 549
          |||||||||| |||||||||||||||||| |||||||| ||||||| |
Db    368 TCTCAGGCAATACTGTCATTTGGGATAATGAAACACCTATTTGTGAGA 415
```

RESULT 9
BABCOREA
LOCUS BABCOREA 945 bp mRNA linear PRI 07-MAY-1996
DEFINITION Papio cynocephalus complement receptor mRNA, partial cds.
ACCESSION L77978
VERSION L77978.1 GI:1301610
KEYWORDS complement C3b; complement receptor; glycophosphatidylinositol-linked protein.
SOURCE Papio cynocephalus cDNA to mRNA.
ORGANISM Papio cynocephalus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Papio.
REFERENCE 1 (bases 1 to 945)
AUTHORS Birmingham,D.J., Logar,C.M., Shen,X.-P. and Chen,W.
TITLE The baboon erythrocyte complement receptor is a glycophosphatidyl inositol-linked protein encoded by a homologue of the human CR1-like genetic element
JOURNAL Unpublished (1996)
FEATURES Location/Qualifiers
source 1. .945
/organism="Papio cynocephalus"
/db_xref="taxon:9556"
/cell_type="lymphocyte"
/tissue_type="bone marrow"
/dev_stage="adult"
mRNA <1. .>945
CDS <1. .>945
/note="homologue of human CR1-like genetic element"
/codon_start=1
/product="complement receptor"
/protein_id="AAA99005.1"
/db_xref="GI:1301611"
/translation="QCNAPEQLPFARPTELIDESEFSIGTHLKYECRPGYYGRPFSII
CLKNSVWTSAKDRCIRKSCRNPRDPVNGMVHVIKDIQFGSQINYSCTEGHRLIGSSSA
TCIISGNTVIWDNETPICEKISCGLPPTIDNGDFFSANKEYFHYGSVVTYRCNLGSGG
RKLFELVGEPSIYCTSNEDQVGIWSGPAPQCIIPNKCTPPNVENGILVSVNRSLFSLN
EVVEFRCQPGFVMKGPRRVQCQALNKWEPELPSCSRVCQPPPDVLHGERTQRDKDIFQ
TGQEVFYICEPSYDLRGAASLRCTPQGDWSPAAPRCEVK"
BASE COUNT 252 a 224 c 228 g 239 t 2 others
ORIGIN

Query Match 50.5%; Score 316.8; DB 9; Length 945;
Best Local Similarity 91.5%; Pred. No. 2.9e-82;
Matches 333; Conservative 2; Mismatches 29; Indels 0; Gaps 0;

```
Qy    186 caatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgacttt 245
          |||||||||| |||||||  |||||||||||||||||||||: | |||| |||||| | |
Db      1 CAATGCAATGCCCCGGAACAGCTTCCATTTGCCAGGCCTACYGAACTAATTGATGAGTCT 60

Qy    246 gagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaagaccg 305
          |||||| |||||||||||| ||||||| |||||||| ||||||||||||   ||||||||:
Db     61 GAGTTTTCCATTGGGACACATCTGAAGTATGAATGTCGCCCTGGTTATTATGGAAGACCR 120

Qy    306 ttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaacgt 365
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||| ||||
Db    121 TTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACAAGTGCTAAGGACAGGTGCATACGT 180

Qy    366 aaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagacatc 425
          |||||||||||||||||  ||||||||||||||||||||  ||||||||||||||||||
Db    181 AAATCATGTCGTAATCCTAGAGATCCTGTGAATGGCATGGTGCATGTGATCAAAGACATC 240

Qy    426 cagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcctcg 485
          |||||||||||||||||| ||||||||| || ||||| ||||||||||||||||||||
Db    241 CAGTTCGGATCCCAAATTAATTATTCTTGTACTGAAGGACACCGACTCATTGGTTCCTCA 300

Qy    486 tctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtttgt 545
          |||||||||||||||||||||||||||| |||||||||||||||||| |||||||| |||||
Db    301 TCTGCCACATGCATCATCTCAGGCAATACTGTCATTTGGGATAATGAAACACCTATTTGT 360

Qy    546 gaca 549
          || |
Db    361 GAGA 364
```

RESULT 10
G25967/c
LOCUS G25967 389 bp DNA linear STS 02-JUN-1996
DEFINITION human STS EST126198, sequence tagged site.
ACCESSION G25967
VERSION G25967.1 GI:1348199
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE human STSs derived from sequences in dbEST and the Unigene collection.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 389)
AUTHORS Hudson,T.
TITLE Whitehead Institute/MIT Center for Genome Research; Physically Mapped STSs
JOURNAL Unpublished
COMMENT
Contact: Thomas Hudson
Whitehead Institute/MIT Center for Genome Research
Whitehead Institute for Biomedical Research
9 Cambridge Center, Cambridge MA 02142 USA
Tel: 617 252 1900
Fax: 617 252 1902
Email: thudson@genome.wi.mit.edu

Primer A: CCAGAGAAAAATTAGAATGTATCGG
Primer B: CTCGTCTGCCACATGCATC
STS size: 125
PCR Profile:
Presoak:
Denaturation:
Annealing: 56 degrees C
Polymerization:
PCR Cycles: 35
Thermal Cycler:
Protocol:
Template: 10 ng
Primer: each 5 pM
dNTPs: each 4 nM
Taq Polymerase: 0.025 units/ul
Total Vol: 20 ul

Buffer:

```
            MgCl2: 1.5 mM
            KCl: 50 mM
            Tris-HCL: 10 mM
            pH: 9.3

         Derived from dbEST (genbank accession T66823).
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="890.2 cR from top of Chr1 linkage group"
     STS             25. .149
     primer_bind     25. .49
     primer_bind     complement(131. .149)
BASE COUNT      114 a     67 c     92 g    112 t      4 others
ORIGIN


  Query Match             48.9%;  Score 306.4;  DB 11;  Length 389;
  Best Local Similarity   95.8%;  Pred. No. 2.8e-79;
  Matches  345;  Conservative    0;  Mismatches   10;  Indels    5;  Gaps    3;

Qy     267 ctgaactatgaatgc---cgccctggttattccggaagaccgtttt-ctatcatctg-cc 321
           ||||||||||||||||   | ||  | |  | ||||||||||||||| |||||||||| ||
Db     369 CTGAACTATGAATGCCCGCCCCTGGTTANTCCCGGAAGACCGTTTTCCTATCATCTGCCC 310

Qy     322 taaaaaactcagtctggacaagtgctaaggacaagtgcaaacgtaaatcatgtcgtaatc 381
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     309 TAAAAAACTCAGTCTGGACAAGTGCTAAGGACAAGTGCAAACGTAAATCATGTCGTAANC 250

Qy     382 ctccagatcctgtgaatggcatggcacatgtgatcaaagacatccagttcggatcccaaa 441
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     249 CTCCAGANCCTGTGAATGGCATGGCACATGTGATCAAAGACATCCAGTTCGGATCCCAAA 190

Qy     442 ttaaatattcttgtcctaaaggataccgactcattggttcctcgtctgccacatgcatca 501
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     189 TTAAATANTCTTGTCCTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCA 130

Qy     502 tctcaggcaacactgtcatttgggataataaaacacctgtttgtgacagtgagttgaaat 561
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129 TCTCAGGCAACACTGTCATTTGGGATAATAAAACACCTGTTTGTGACAGTGAGTTGAAAT 70

Qy     562 atgcattcctatttcttttaccgatacattctaatttttctctggaataataaaaatctt 621
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69 ATGCATTCCTATTTCTTTTACCGATACATTCTAATTTTTCTCTGGAATAATAAAAATCTT 10


RESULT  11
G27827/c
LOCUS       G27827                   389 bp    DNA     linear   STS 29-JUN-1996
DEFINITION  human STS SHGC-33387, sequence tagged site.
ACCESSION   G27827
VERSION     G27827.1  GI:1396546
KEYWORDS    STS; STS sequence; primer; sequence tagged site.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 389)
  AUTHORS   Myers,R.M.
  JOURNAL   Unpublished
COMMENT
            Contact: Richard M. Myers
            Stanford Human Genome Center (SHGC)
            Stanford University School of Medicine
            Department of Genetics, M-344, Stanford, CA 94305, USA
            Tel: 4157259687
            Fax: 4157259689
            Email: myers@shgc.stanford.edu

            Primer A: CCAGAGAAAAATTAGAATGTATCGG
            Primer B: CTCGTCTGCCACATGCATC
            STS size: 125
            PCR Profile:
                    Initial incubation: 94 degrees C for 90 seconds

                    Denaturation:       94 degrees C for 15 seconds
                    Annealing:          62 degrees C for 23 seconds
                    Polymerization:     72 degrees C for 30 seconds
                    PCR Cycles:         30
                    Thermal Cycler:     Perkin Elmer 9600
            Protocol:
                    Template:       25 ng
                    Primer:         each 1 uM
                    dNTPs:          each 200 uM
                    Taq Polymerase: 0.05 units/ul
                    Total Vol:      10 ul

            Buffer:
                    MgCl2:          2.5 mM
                    KCl:            50  mM
                    Tris-HCl:       20  mM
                    pH:             8.3

            Prepared with primer pairs provided by Sandoz, derived from T66823
            -- Washington University/Merck EST sequence.
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="1"
     STS             25. .149
     primer_bind     25. .49
     primer_bind     complement(131. .149)
BASE COUNT      114 a     67 c     92 g    112 t      4 others
ORIGIN


  Query Match             48.9%;  Score 306.4;  DB 11;  Length 389;
  Best Local Similarity   95.8%;  Pred. No. 2.8e-79;
  Matches  345;  Conservative    0;  Mismatches   10;  Indels    5;  Gaps    3;

Qy     267 ctgaactatgaatgc---cgccctggttattccggaagaccgtttt-ctatcatctg-cc 321
           ||||||||||||||||   | ||  | |  | ||||||||||||||| |||||||||| ||
Db     369 CTGAACTATGAATGCCCGCCCCTGGTTANTCCCGGAAGACCGTTTTCCTATCATCTGCCC 310

Qy     322 taaaaaactcagtctggacaagtgctaaggacaagtgcaaacgtaaatcatgtcgtaatc 381
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     309 TAAAAAACTCAGTCTGGACAAGTGCTAAGGACAAGTGCAAACGTAAATCATGTCGTAANC 250

Qy     382 ctccagatcctgtgaatggcatggcacatgtgatcaaagacatccagttcggatcccaaa 441
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     249 CTCCAGANCCTGTGAATGGCATGGCACATGTGATCAAAGACATCCAGTTCGGATCCCAAA 190

Qy     442 ttaaatattcttgtcctaaaggataccgactcattggttcctcgtctgccacatgcatca 501
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     189 TTAAATANTCTTGTCCTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCA 130

Qy     502 tctcaggcaacactgtcatttgggataataaaacacctgtttgtgacagtgagttgaaat 561
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129 TCTCAGGCAACACTGTCATTTGGGATAATAAAACACCTGTTTGTGACAGTGAGTTGAAAT 70

Qy     562 atgcattcctatttcttttaccgatacattctaatttttctctggaataataaaaatctt 621
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      69 ATGCATTCCTATTTCTTTTACCGATACATTCTAATTTTTCTCTGGAATAATAAAAATCTT 10


RESULT  12
CHPCR1Y
LOCUS       CHPCR1Y                  1731 bp    mRNA    linear   PRI 12-OCT-1994
DEFINITION  Pan troglodytes mRNA sequence, 3'end of ORF.
ACCESSION   L24922
VERSION     L24922.1  GI:557728
KEYWORDS    .
SOURCE      Pan troglodytes cDNA to mRNA.
  ORGANISM  Pan troglodytes
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 1731)
  AUTHORS   Birmingham,D.J., Shen,X.P., Hourcade,D., Nickells,M.W. and
            Atkinson,J.P.
  TITLE     Primary sequence of an alternatively spliced form of CR1. Candidate
            for the 75,000 M(r) complement receptor expressed on chimpanzee
            erythrocytes
  JOURNAL   J. Immunol. 153 (2), 691-700 (1994)
  MEDLINE   94292799
FEATURES             Location/Qualifiers
     source          1. .1731
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
                     /cell_line="EBV transformed"
                     /cell_type="B cell"
     CDS             <1. .1571
                     /note="Possible homologue to human CR1-like genomic
                     element"
                     /codon_start=3
                     /protein_id="AAA50460.1"
                     /db_xref="GI:557729"
                     /translation="DEFEFPIGTYLNYECRPGYYGRPFSIICLKNSVWTSAEDKCKRK
                     SCRNPPDPVNGMVHVIKDIQFGSQIKYSCTKGYRLIGSSSATCIISGNTVIWDNKTPV
                     CDRIICGLPPTIANGDFTSISREYFHYASVVTYHCNLGSGGKKVFELVGEPSIYCTSK
                     DDQVGIWSGPAPQCIIPNKCTPPNVENGILVSDNRSLFSLNEVVEFRCQPGFVMKGPR
                     HVHCQALNKWEPELPSCSRVCQPPPDVLHGERTQRDKDNFSPGEEVYYSCEPGYDLRG
                     STYLHCTPQGDWSPEAPRCEVKSCDDFLGQLPNGRVLFPLNLQLGAKVDFVCDEGFQL
                     KGSSASYCVLAGMESLWNSSVPVCERKSCETPPVPVNGMVHVITDIHVGSRINXSCIT
                     GFRLKGRSASHCVLAGMKALWNSSVPVCEQIFCPNPPAILNGRHTGTPLGDIPYGKEV
                     SYTCDPHPDRGMTFNLIGESTIRCTSDPHGNGVWSSPAPRCELPVGAGSHDALIVGTL
                     LGTIIFILIIIFLY"
     3'UTR           1572. .>1731
BASE COUNT      469 a    395 c    403 g    463 t      1 others
ORIGIN


  Query Match            48.4%;  Score 303.2;  DB 9;  Length 1731;
  Best Local Similarity  97.5%;  Pred. No. 3.3e-78;
  Matches  308;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    235 ctgatgactttgagtttcccattgggacatatctgaactatgaatgccgccctggttatt 294
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CTGATGAGTTTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATT 60

Qy    295 ccggaagaccgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggaca 354
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     61 ATGGAAGACCATTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACAAGTGCTGAGGACA 120

Qy    355 agtgcaaacgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtga 414
          |||||||||||||||||||||||||||||||||||||||||||||||||||||  |||||||
Db    121 AGTGCAAACGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATGGTGCATGTGA 180

Qy    415 tcaaagacatccagttcggatcccaaattaaatattcttgtcctaaaggataccgactca 474
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    181 TCAAAGACATCCAGTTCGGATCCCAAATTAAATATTCTTGTACTAAAGGATACCGACTCA 240

Qy    475 ttggttcctcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaa 534
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTGGTTCCTCGTCTGCCACATGCATCATCTCAGGCAACACTGTCATTTGGGATAATAAAA 300

Qy    535 cacctgtttgtgacag 550
          ||||||||||||||||
Db    301 CACCTGTTTGTGACAG 316


RESULT  13
A86593
LOCUS       A86593                   591 bp    DNA     linear   PAT 21-JAN-2000
DEFINITION  Sequence 2 from Patent WO9839433.
ACCESSION   A86593
VERSION     A86593.1  GI:6735168
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Smith,R.A. and Cox,V.F.
  TITLE     COMPLEMENT RECEPTOR TYPE 1 (CR1)-LIKE SEQUENCES
  JOURNAL   Patent: WO 9839433-A 2 11-SEP-1998;
            SMITH RICHARD ANTONY GODWIN (GB); ADPROTECH PLC (GB)
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      132 a    159 c    148 g    152 t
ORIGIN


  Query Match            42.6%;  Score 267.4;  DB 6;  Length 591;
  Best Local Similarity  83.3%;  Pred. No. 9.4e-68;
  Matches  304;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy    186 caatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgacttt 245
          || ||||| |  |||||||||||| || || ||  | || |||||||| |||||||| |||
Db      4 CAGTGCAACGCTCCGGAATGGCTGCCGTTCGCGCGCCCGACCAACCTGACTGATGAATTT 63

Qy    246 gagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaagaccg 305
          ||||| || || || || || |||||||| |||||||||||| ||||||  |||  | |||
Db     64 GAGTTCCCGATCGGTACCTACCTGAACTACGAATGCCGCCCGGGTTATAGCGGCCGCCCG 123

Qy    306 ttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaacgt 365
          |||||||||||||||| |||||||| |||||||| ||||||||||   |||  ||||
Db    124 TTTTCTATCATCTGCCTGAAAAACTCTGTCTGGACTGGTGCTAAGGACCGTTGCCGACGT 183

Qy    366 aaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagacatc 425
          ||||| |||||||||| |||||||| || || |||||||  |||||||||||| ||||
Db    184 AAATCTTGTCGTAATCCGCCAGATCCGGTTAACGGCATGGTGCATGTGATCAAAGGCATC 243

Qy    426 cagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcctcg 485
          |||||||| |||||||||||||||||||| ||||||| ||||| || ||||||||||||
Db    244 CAGTTCGGTTCCCAAATTAAATATTCTTGTACTAAAGGTTACCGTCTGATTGGTTCCTCC 303

Qy    486 tctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtttgt 545
          || |||||||||||||| ||  | ||||||||||||||||||| |||||||  |||||
Db    304 AGCGCTACATGCATCATCTCTGGTGATACTGTCATTTGGGATAATGAAACACCGATTTGT 363

Qy    546 gacag 550
          |||| |
Db    364 GACCG 368


RESULT  14
A86601
LOCUS       A86601                   591 bp    DNA     linear   PAT 21-JAN-2000
DEFINITION  Sequence 10 from Patent WO9839433.
ACCESSION   A86601
VERSION     A86601.1  GI:6735175
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Smith,R.A. and Cox,V.F.
  TITLE     COMPLEMENT RECEPTOR TYPE 1 (CR1)-LIKE SEQUENCES
  JOURNAL   Patent: WO 9839433-A 10 11-SEP-1998;
            SMITH RICHARD ANTONY GODWIN (GB); ADPROTECH PLC (GB)
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      127 a    159 c    151 g    154 t
ORIGIN


  Query Match            42.6%;  Score 267.4;  DB 6;  Length 591;
```

```
  Best Local Similarity   83.3%;  Pred. No. 9.4e-68;
  Matches  304;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy    186 caatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgacttt 245
          || ||||| |  |||||||||||| || || ||  | || |||||||| ||||||| |||
Db      4 CAGTGCAACGCTCCGGAATGGCTGCCGTTCGCGCGCCCGACCAACCTGACTGATGAATTT 63

Qy    246 gagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaagaccg 305
          ||||| || || || || || |||||||| |||||||||| |||||| ||| | |||
Db     64 GAGTTCCCGATCGGTACCTACCTGAACTACGAATGCCGCCCGGGTTATAGCGGCCGCCCG 123

Qy    306 ttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaacgt 365
          |||||||||||||||| |||||||| |||||||| |||||||||||   ||| ||||
Db    124 TTTTCTATCATCTGCCTGAAAAACTCTGTCTGGACTGGTGCTAAGGACCGTTGCCGACGT 183

Qy    366 aaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagacatc 425
          ||||| |||||||||| |||||||| || || |||||||  |||||||||||| ||||
Db    184 AAATCTTGTCGTAATCCGCCAGATCCGGTTAACGGCATGGTGCATGTGATCAAAGGCATC 243

Qy    426 cagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcctcg 485
          |||||||| |||||||||||||||||||| ||||||| ||||| || ||||||||||||
Db    244 CAGTTCGGTTCCCAAATTAAATATTCTTGTACTAAAGGTTACCGTCTGATTGGTTCCTCC 303

Qy    486 tctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtttgt 545
             || |||||||||||||| ||  | |||||||||||||||||| |||||||  |||||
Db    304 AGCGCTACATGCATCATCTCTGGTGATACTGTCATTTGGGATAATGAAACACCGATTTGT 363

Qy    546 gacag 550
          |||| |
Db    364 GACCG 368
```

```
RESULT  15
A86603
LOCUS       A86603                   591 bp    DNA     linear   PAT 21-JAN-2000
DEFINITION  Sequence 12 from Patent WO9839433.
ACCESSION   A86603
VERSION     A86603.1  GI:6735176
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Smith,R.A. and Cox,V.F.
  TITLE     COMPLEMENT RECEPTOR TYPE 1 (CR1)-LIKE SEQUENCES
  JOURNAL   Patent: WO 9839433-A 12 11-SEP-1998;
            SMITH RICHARD ANTONY GODWIN (GB); ADPROTECH PLC (GB)
FEATURES             Location/Qualifiers
     source          1. .591
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      134 a    158 c    146 g    153 t
ORIGIN
```

```
  Query Match             42.6%;  Score 267.4;  DB 6;  Length 591;
  Best Local Similarity   83.3%;  Pred. No. 9.4e-68;
  Matches  304;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy    186 caatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgacttt 245
          || ||||| |  |||||||||||| || || ||  | || |||||||| ||||||| |||
Db      4 CAGTGCAACGCTCCGGAATGGCTGCCGTTCGCGCGCCCGACCAACCTGACTGATGAATTT 63

Qy    246 gagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaagaccg 305
          ||||| || || || || || |||||||| |||||||||| |||||| ||| | |||
Db     64 GAGTTCCCGATCGGTACCTACCTGAACTACGAATGCCGCCCGGGTTATAGCGGCCGCCCG 123

Qy    306 ttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaacgt 365
          |||||||||||||||| |||||||| |||||||| |||||||||||   ||| ||||
Db    124 TTTTCTATCATCTGCCTGAAAAACTCTGTCTGGACTGGTGCTAAGGACCGTTGCCGACGT 183

Qy    366 aaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagacatc 425
          ||||| |||||||||||| |||||||| || || |||||||  |||||||||||| ||||
```

```
Db    184 AAATCTTGTCGTAATCCGCCAGATCCGGTTAACGGCATGGTGCATGTGATCAAAGGCATC 243

Qy    426 cagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcctcg 485
          |||||||| |||||||||||||||||||| ||||||| ||||| || ||||||||||||
Db    244 CAGTTCGGTTCCCAAATTAAATATTCTTGTACTAAAGGTTACCGTCTGATTGGTTCCTCC 303

Qy    486 tctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtttgt 545
             || |||||||||||||| ||  | |||||||||||||||||| |||||||  |||||
Db    304 AGCGCTACATGCATCATCTCTGGTGATACTGTCATTTGGGATAATGAAACACCGATTTGT 363

Qy    546 gacag 550
          |||| |
Db    364 GACCG 368
```

```
Search completed: October  9, 2002, 17:10:18
Job time: 4235 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 9, 2002, 16:37:29 ; Search time 176.28 Seconds
(without alignments)
6106.792 Million cell updates/sec

Title: US-10-031-904-30
Perfect score: 627
Sequence: 1 cggactcagaagggacttcc.....ataataaaaatcttaaccga 627

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	627	100.0	627	22	AAF58602	Human RECAP polynu
	2	352	56.1	6951	10	AAN91477	CR1 protein DNA.
	3	352	56.1	6951	12	AAQ11642	Entire human compl
	4	352	56.1	6951	20	AAZ38150	Human C3b/C4b rece
	5	352	56.1	7313	22	AAI58380	Human polynucleoti
	6	351.8	56.1	7821	23	AAS64474	DNA encoding novel
	7	348	55.5	6951	14	AAQ41867	CR1 coding region.
	8	341	54.4	7028	22	ABA09026	Human CR1 protein
	9	341	54.4	7028	22	AAI60166	Human polynucleoti
	10	273.2	43.6	9038	23	AAS64290	DNA encoding novel
	11	269.4	43.0	2796	23	AAS64473	DNA encoding novel
	12	267.4	42.6	591	19	AAV53262	Complement recepto
	13	267.4	42.6	591	19	AAV53269	Complement recepto
	14	267.4	42.6	591	19	AAV53270	Complement recepto
	15	267.4	42.6	591	19	AAV53271	Complement recepto
	16	267.4	42.6	591	19	AAV53272	Complement recepto
	17	267.4	42.6	591	19	AAV53273	Complement recepto
	18	247.8	39.5	2798	23	AAS64289	DNA encoding novel
	19	218.6	34.9	5420	12	AAQ11643	Partial human comp
	20	218.6	34.9	5420	20	AAZ38151	Human C3b/C4b rece
	21	156.4	24.9	1477	18	AAT90306	Human MCP cDNA. H
	22	156.4	24.9	3302	22	AAH34933	Human colon cancer
	23	156.4	24.9	3308	21	AAF18270	Lung cancer associ
	24	153.6	24.5	396	21	AAC04275	Human secreted pro
c	25	124.6	19.9	125	19	AAX11827	Human biallelic po
	26	120	19.1	1437	23	AAS64286	DNA encoding novel
	27	120	19.1	1437	23	AAS64470	DNA encoding novel
	28	117.6	18.8	1530	16	AAQ99106	Human MCP (CD46) f
	29	117.6	18.8	1530	16	AAT03339	Human CD46 cDNA.
	30	117.6	18.8	1546	12	AAQ10864	Sequence encoding
	31	116.6	18.6	1878	17	AAT46065	Membrane co-factor
	32	103.4	16.5	1244	12	AAQ14919	CD46 clone pm5.8.
	33	103.4	16.5	1659	12	AAQ14915	CD46 clone pm5.1.
	34	103.4	16.5	1991	12	AAQ14916	CD46 clone pm5.6.
c	35	103.4	16.5	9888	24	ABL33241	Human immune syste
	36	88.8	14.2	1848	22	AAD20355	CAB2 protein encod
	37	88.8	14.2	1848	22	AAD08170	CAB-2 DNA. Uniden
c	38	88	14.0	7821	23	AAS64474	DNA encoding novel
	39	87.4	13.9	945	17	AAT17599	CD46 construct del
	40	87.4	13.9	1134	17	AAT17598	CD46 construct sub
	41	87.4	13.9	1134	17	AAT17595	CD46 wild-type cDN
	42	87.4	13.9	1134	17	AAT17596	CD46 construct sub
	43	87.4	13.9	1134	17	AAT17597	CD46 construct sub
	44	84.8	13.5	9888	24	ABL33240	Human immune syste
	45	79.4	12.7	1104	15	AAQ58894	Sequence encoding

ALIGNMENTS

RESULT 1
AAF58602
ID AAF58602 standard; cDNA; 627 BP.
XX
AC AAF58602;
XX
DT 24-APR-2001 (first entry)
XX
DE Human RECAP polynucleotide, SEQ ID NO: 30.
XX
KW Human; RECAP; receptors and associated proteins; cerebroprotective;
KW nootropic; neuroprotective; anticonvulsant; antiparkinsonian; anti-HIV;
KW antidiabetic; immunostimulant; immunomodulator; antiinflammatory;
KW antithyroid; immunosuppressive; nephrotropic; antigout; thyromimetic;
KW cytostatic; antibacterial; virucide; fungicide; protozoacide;
KW antiarteriosclerotic; hepatotropic; gene therapy; infection; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200107612-A2.
XX
PD 01-FEB-2001.
XX
PF 21-JUL-2000; 2000WO-US20035.
XX
PR 21-JUL-1999; 99US-0145232.
PR 07-OCT-1999; 99US-0158578.
PR 12-NOV-1999; 99US-0165192.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Au-Young J, Bandman O, Tang YT, Yue H, Azimzai Y, Burford N;

```
PI   Baughn MR,  Lu DAM,  Hillman JL,  Patterson C,  Lal P;
XX
DR   WPI; 2001-168554/17.
DR   P-PSDB; AAB68878.
XX
PT   Novel receptors and associated proteins for diagnosis and treatment of
PT   neurological disorders, immunological disorders including autoimmune/
PT   inflammatory disorders and cell proliferative disorders such as cancer
PT   -
XX
PS   Claim 5; Page 120; 128pp; English.
XX
CC   The present sequence encodes a human RECAP (receptors and associated
CC   proteins) polypeptide. RECAP polynucleotides and polypeptides are useful
CC   in the diagnosis, treatment and prevention of neurological disorders
CC   such as stroke, Alzheimer's disease, Pick's disease, Huntington's
CC   disease, dementia, Parkinson's disease, Downs's syndrome, amyotrophic
CC   lateral sclerosis, multiple sclerosis, bacterial and viral meningitis,
CC   CJD (Creutzfeldt-Jakob disease), GSS (Gerstmann -Straussler-Scheinker
CC   syndrome); immunological disorders, including autoimmune/inflammatory
CC   disorders such as AIDS, DiGeorge's syndrome, severe combined
CC   immunodeficiency disease (SCID), Chediak-Higashi syndrome, Cushing's
CC   disease, Addison's disease, autoimmune thyroiditis, Crohn's disease,
CC   diabetes mellitus, Good pasture's syndrome, gout, Grave's diseases,
CC   Hashimoto's thyroiditis, Sjogren's syndrome, Werner's syndrome, viral,
CC   bacterial, fungal, parasitic, protozoal, and helminthic infections; and
CC   cell proliferation disorders such as arteriosclerosis, atherosclerosis,
CC   cirrhosis, hepatitis and cancer.
XX
SQ   Sequence 627 BP; 142 A; 173 C; 128 G; 184 T; 0 other;


  Query Match             100.0%;  Score 627;  DB 22;  Length 627;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-181;
  Matches  627;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 cggactcagaagggacttccctgctcggctggctttcggtttctctgctcacctccggat 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 cggactcagaagggacttccctgctcggctggctttcggtttctctgctcacctccggat 60

Qy      61 aaatcacggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtccctttc 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 aaatcacggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtccctttc 120

Qy     121 cttcccggcgctttcctgggttgcttctggcggccctggtgttgctgctgtcctccttct 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 cttcccggcgctttcctgggttgcttctggcggccctggtgttgctgctgtcctccttct 180

Qy     181 ccgatcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatg 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ccgatcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatg 240

Qy     241 actttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaa 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 actttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaa 300

Qy     301 gaccgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgca 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 gaccgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgca 360

Qy     361 aacgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaag 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 aacgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaag 420

Qy     421 acatccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggtt 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 acatccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggtt 480

Qy     481 cctcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctg 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 cctcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctg 540

Qy     541 tttgtgacagtgagttgaaatatgcattcctatttcttttaccgatacattctaattttt 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 tttgtgacagtgagttgaaatatgcattcctatttcttttaccgatacattctaattttt 600

Qy     601 ctctggaataataaaaatcttaaccga 627
           |||||||||||||||||||||||||||
Db     601 ctctggaataataaaaatcttaaccga 627


RESULT   2
AAN91477
ID   AAN91477 standard; DNA; 6951 BP.
XX
AC   AAN91477;
XX
DT   22-FEB-1990  (first entry)
XX
DE   CR1 protein DNA.
XX
KW   Complement; cofactor.
XX
OS   Homo sapiens (human).
XX
FH   Key             Location/Qualifiers
FT   sig_peptide     28..147
FT                   /*tag=  a
FT   CDS             1..6951
FT                   /*tag=  b
FT   CDS             28..6147
FT                   /*tag=  c
FT                   /note="CR1 protein"
FT   CDS             28..1533
FT                   /*tag=  d
FT                   /note="CR1 protein"
XX
PN   WO8909220-A.
XX
PD   05-OCT-1989.
XX
PF   31-MAR-1989;   89WO-US01358.
XX
PR   01-APR-1988;   88US-0176532.
XX
PA   (TCEL-) T CELL SCIENCES INC.
PA   (UYJO ) THE JOHNS HOPKINS UNIVERSITY.
PA   (BRIG ) THE BRIGHAM AND WOMENS HOSPITAL.
XX
PI   Fearon DT,  Klickstein LB,  Wong W,  Carson G,  Concino MF,  Makrides SC;
XX
DR   WPI; 1989-309498/42.
DR   P-PSDB; AAP92219.
XX
PT   New nucleic acid sequences encoding new CR1 protein - and its fragment,
PT   for diagnosis and control of complement-related immune defects,
PT   inflammation, myocardial infarct, etc
XX
PS   Claim 1; fig. 1; 191pp; English.
XX
CC   This is full-length CR1 protein, and shortened forms are new, lacking
CC   the transmembrane region. The proteins and fragments bind C3b and/or
CC   C4b, have cofactor activity and inhibit C3 and C5 convertase activity.
CC   They are useful in diagnosing and treating immune disorders, and prevent
CC   perfusion injury.
XX
SQ   Sequence 6951 BP; 1802 A; 1680 C; 1660 G; 1809 T; 0 other;


  Query Match             56.1%;  Score 352;  DB 10;  Length 6951;
  Best Local Similarity   84.0%;  Pred. No. 1.3e-96;
  Matches  410;  Conservative    0;  Mismatches   75;  Indels    3;  Gaps    1;

Qy      66 acggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtccctttccttcc 125
           | |||| |  |  | ||    |   ||   || ||| | |    |   |||  ||   ||
Db      28 atgggggcctcttctccaagaagcccggagcctgtcgggccgccggcgcccggtctcccc 87
```

```
Qy    126 cggcgctttcctgggttgcttctggcggccctggtgttgct---gctgtcctccttctcc 182
          |||     || |  || |||||||   ||||| ||||   |||| |      ||
Db     88 ttctgctgcggaggatccctgctggcggttgtggtgctgcttgcgctgccggtggcctgg 147

Qy    183 gatcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgac 242
          | ||||||||||| ||| ||||||||||||||||||||||||||||||||||||||||
Db    148 ggtcaatgcaatgccccagaatggcttccatttgccaggcctaccaacctaactgatgag 207

Qy    243 tttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaaaga 302
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    208 tttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaaaga 267

Qy    303 ccgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaa 362
          |||||||||||||||||||||||||||||||||||||  |||||||||||| ||||| |
Db    268 ccgttttctatcatctgcctaaaaaactcagtctggactggtgctaaggacaggtgcaga 327

Qy    363 cgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagac 422
          |||||||||||||||||||||||||||||||||||||||||||  |||||||||||| |
Db    328 cgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggtgcatgtgatcaaaggc 387

Qy    423 atccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcc 482
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    388 atccagttcggatcccaaattaaatattcttgtactaaaggataccgactcattggttcc 447

Qy    483 tcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtt 542
          |||||||||||||||||||||||||  | ||||||||||||||||| ||||||||  ||
Db    448 tcgtctgccacatgcatcatctcaggtgatactgtcatttgggataatgaaacacctatt 507

Qy    543 tgtgacag 550
          ||||||||
Db    508 tgtgacag 515
```

```
RESULT   3
AAQ11642
ID   AAQ11642 standard; DNA; 6951 BP.
XX
AC   AAQ11642;
XX
DT   25-JUN-1991  (first entry)
XX
DE   Entire human complement type 1 receptor coding region.
XX
KW   complement system; C3b/C4b receptor; CR1; allergic reaction;
KW   immune response; clone lambda T109.1; ss.
XX
     Homo sapiens.

     Key             Location/Qualifiers
FT   sig_peptide     28..147
FT                   /*tag=  a
FT                   /note= "putative"
FT   mat_peptide     148..6144
FT                   /*tag=  b
FT                   /product= C3b/C4b receptor
XX
PN   WO9105047-A.
XX
PD   18-APR-1991.
XX
PF   25-SEP-1990;   90WO-US05454.
XX
PR   26-SEP-1989;   89US-0412745.
PR   26-SEP-1990;   90US-0912349.
XX
PA   (TCEL-) T CELL SCI INC.
PA   (UYJO ) JOHNS HOPKINS UNIVERSITY.
PA   (BRIG-) BRIGHAM AND WOMEN'S HOSPITAL.
XX
PI   Fearon DT,  Klickstein LB,  Wong WW,  Carson GR,  Hoh M,  Concino MF;
PI   Makrides SC,  Marsh HC;
XX
DR   WPI; 1991-132854/18.
DR   P-PSDB; AAR11810.
XX
PT   Human complement receptor type 1 gene, encoded proteins and
PT   fragments - for treatment of immune disorders, myocardial infarct,
PT   damage due to inflammmation and in treatment of thrombosis
XX
PS   Claim 1; Fig 1; 234pp; English.
XX
CC   This sequence is a composite of sequences isolated as lambda clones
CC   H10.3, T109.1, H3.1 and H7.1. The clones were present in the
CC   specifically primed lambda gt11 cDNA library (Lambda HH) which was
CC   prepared with cDNA synthesised from poly(A)+ RNA from DMSO induced
CC   HL-60 cells. The library was screened using probes CR1-1, CR-2 and
CC   CR1-4 (see Wong,WW et al.,1985, Proc. Nat. Acad. Sci. USA. 82:7711)
CC   and probe CR1-18 (corresponding to nucleotides 101-352 of this
CC   sequence). There are four direct, long homologous repeats of 450bp,
CC   each comprising 7 short consensus repeats. Nucleotides 28-1533 are
CC   also claimed separately.
CC   See also AAQ11643.
XX
SQ   Sequence 6951 BP; 1802 A; 1681 C; 1659 G; 1809 T; 0 other;
```

```
  Query Match             56.1%;  Score 352;  DB 12;  Length 6951;
  Best Local Similarity   84.0%;  Pred. No. 1.3e-96;
  Matches  410;  Conservative    0;  Mismatches   75;  Indels    3;  Gaps    1;

Qy     66 acggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtccctttccttcc 125
          | ||||  |  | ||    |   ||   || ||| | |    |   |||  ||   ||
Db     28 atgggggcctcttctccaagaagcccggagcctgtcgggccgccggcgcccggtctcccc 87

Qy    126 cggcgctttcctgggttgcttctggcggccctggtgttgct---gctgtcctccttctcc 182
          |||     || |  || |||||||   ||||| ||||   |||| |      ||
Db     88 ttctgctgcggaggatccctgctggcggttgtggtgctgcttgcgctgccggtggcctgg 147

Qy    183 gatcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgac 242
          | ||||||||||| ||| ||||||||||||||||||||||||||||||||||||||||
Db    148 ggtcaatgcaatgccccagaatggcttccatttgccaggcctaccaacctaactgatgag 207

Qy    243 tttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaaaga 302
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    208 tttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaaaga 267

Qy    303 ccgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaa 362
          |||||||||||||||||||||||||||||||||||||  |||||||||||| ||||| |
Db    268 ccgttttctatcatctgcctaaaaaactcagtctggactggtgctaaggacaggtgcaga 327

Qy    363 cgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagac 422
          |||||||||||||||||||||||||||||||||||||||||||  |||||||||||| |
Db    328 cgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggtgcatgtgatcaaaggc 387

Qy    423 atccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcc 482
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    388 atccagttcggatcccaaattaaatattcttgtactaaaggataccgactcattggttcc 447

Qy    483 tcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtt 542
          |||||||||||||||||||||||||  | ||||||||||||||||| ||||||||  ||
Db    448 tcgtctgccacatgcatcatctcaggtgatactgtcatttgggataatgaaacacctatt 507

Qy    543 tgtgacag 550
          ||||||||
Db    508 tgtgacag 515
```

```
RESULT   4
AAZ38150
ID   AAZ38150 standard; DNA; 6951 BP.
XX
AC   AAZ38150;
XX
DT   22-FEB-2000  (first entry)
XX
```

```
DE   Human C3b/C4b receptor (CR1) protein encoding DNA.
XX
KW   C3B/C4B receptor; CR1 protein; cell-surface protein; erythrocyte; human;
KW   complement regulatory activity; complement pathway enzyme; tissue damage;
KW   reperfusion injury; Arthus reaction; myocardial infarct; inflammation;
KW   heart condition; autoimmune disorder; diagnostic; ss.
XX
OS   Homo sapiens.
XX
PN   US5981481-A.
XX
PD   09-NOV-1999.
XX
PF   06-JUN-1995;   95US-0470652.
XX
PR   03-APR-1989;   89US-0332865.
PR   06-DEC-1974;   74US-0350238.
PR   24-FEB-1993;   93US-0026134.
PR   01-APR-1988;   88US-0176532.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
PA   (AVAN-) AVANT IMMUNOTHERAPEUTICS INC.
XX
PI   Concino MF,  Wong WW,  Makrides SC,  Klickstein LB,  Fearon DT,  Ip SH;
PI   Marsh HC,  Carson GR;
XX
DR   WPI; 1999-633357/54.
DR   P-PSDB; AAY55751.
XX
PT   A human C3B/C4B receptor (CR1) protein having antiinflammatory and
PT   cardiant activity -
XX
PS   Disclosure; Fig 1A-P; 87pp; English.
XX
CC   The invention relates to a human C3B/C4B receptor (CR1) protein. The CR1
CC   protein or fragment is expressed as a cell-surface protein on the surface
CC   of a non-human cell and exhibits a complement regulatory activity of full
CC   -length human CR1 as expressed on erythrocytes. The CR1 function in vivo
CC   may be mediated through the inhibition of complement pathway enzymes. The
CC   soluble CR1 protein exhibits a complement regulatory activity, and this
CC   may be used to prevent reperfusion injury, inhibit Arthus reaction, and
CC   neutrophil mediated tissue damage, and reduce myocardial infarct size,
CC   and inflammation. The CR1 protein and its fragments can also be used in
CC   the treatment of conditions which involve unwanted complement activity,
CC   e.g. shock lung, tissue damage due to burn, or ischemic heart conditions,
CC   and autoimmune disorders. CR1 proteins, analogues, derivatives, and anti
CC   -CR1 antibodies are used in assays, and diagnostics. The present sequence
CC   represents a DNA encoding the human CR1 protein.
XX
SQ   Sequence 6951 BP; 1802 A; 1680 C; 1661 G; 1808 T; 0 other;
```

```
  Query Match             56.1%;  Score 352;  DB 20;  Length 6951;
  Best Local Similarity   84.0%;  Pred. No. 1.3e-96;
  Matches  410;  Conservative    0;  Mismatches   75;  Indels    3;  Gaps    1;

Qy      66 acggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtccctttccttcc 125
           | |||| |  |  | ||    |   ||   || ||| | |    |   |||  ||   ||
Db      28 atgggggcctcttctccaagaagcccggagcctgtcgggccgccggcgcccggtctcccc 87

Qy     126 cggcgctttcctgggttgcttctggcggccctggtgttgct---gctgtcctccttctcc 182
               |||     || |  || ||||||||   ||||| ||||   |||| |      ||
Db      88 ttctgctgcggaggatccctgctggcggttgtggtgctgcttgcgctgccggtggcctgg 147

Qy     183 gatcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgac 242
           | |||||||||||| ||| ||||||||||||||||||||||||||||||||||||||
Db     148 ggtcaatgcaatgccccagaatggcttccatttgccaggcctaccaacctaactgatgag 207

Qy     243 tttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaaga 302
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208 tttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaaga 267

Qy     303 ccgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaa 362
           |||||||||||||||||||||||||||||||||||||||  ||||||||||| ||||| |
Db     268 ccgttttctatcatctgcctaaaaaactcagtctggactggtgctaaggacaggtgcaga 327

Qy     363 cgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagac 422
           |||||||||||||||||||||||||||||||||||||||||  |||||||||||| |
Db     328 cgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggtgcatgtgatcaaaggc 387

Qy     423 atccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcc 482
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     388 atccagttcggatcccaaattaaatattcttgtactaaaggataccgactcattggttcc 447

Qy     483 tcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtt 542
           |||||||||||||||||||||||||  | |||||||||||||||||| ||||||||  ||
Db     448 tcgtctgccacatgcatcatctcaggtgatactgtcatttgggataatgaaacacctatt 507

Qy     543 tgtgacag 550
           ||||||||
Db     508 tgtgacag 515
```

```
RESULT   5
AAI58380
ID   AAI58380 standard; cDNA; 7313 BP.
XX
AC   AAI58380;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 583.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM39224.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 583; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
```

```
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 7313 BP; 1903 A; 1770 C; 1733 G; 1907 T; 0 other;
```

```
  Query Match             56.1%;  Score 352;  DB 22;  Length 7313;
  Best Local Similarity   84.0%;  Pred. No. 1.4e-96;
  Matches  410;  Conservative    0;  Mismatches   75;  Indels    3;  Gaps    1;

      66 acggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtccctttccttcc 125
         | |||| |  |  | ||    |   ||   || ||| | |    |   |||  ||   ||
Db    28 atgggggcctcttctccaagaagcccggagcctgtcgggccgccggcgcccggtctcccc 87

Qy   126 cggcgctttcctgggttgcttctggcggccctggtgttgct---gctgtcctccttctcc 182
            |||     || |. || |||||||   |||| ||||   |||| |       ||
Db    88 ttctgctgcggaggatccctgctggcggttgtggtgctgcttgcgctgccggtggcctgg 147

Qy   183 gatcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgac 242
         | ||||||||||| ||| ||||||||||||||||||||||||||||||||||||||||
Db   148 ggtcaatgcaatgccccagaatggcttccatttgccaggcctaccaacctaactgatgag 207

Qy   243 tttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaaga 302
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   208 tttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaaga 267

Qy   303 ccgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaa 362
         ||||||||||||||||||||||||||||||||||||||  ||||||||||| ||||| |
Db   268 ccgttttctatcatctgcctaaaaaactcagtctggactggtgctaaggacaggtgcaga 327

Qy   363 cgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagac 422
         |||||||||||||||||||||||||||||||||||||  |||||||||||| |
Db   328 cgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggtgcatgtgatcaaaggc 387

Qy   423 atccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcc 482
         |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   388 atccagttcggatcccaaattaaatattcttgtactaaaggataccgactcattggttcc 447

      483 tcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtt 542
          |||||||||||||||||||||||  | ||||||||||||||| ||||||| ||
Db    448 tcgtctgccacatgcatcatctcaggtgatactgtcatttgggataatgaaacacctatt 507

Qy   543 tgtgacag 550
         ||||||||
Db   508 tgtgacag 515
```

```
RESULT   6
AAS64474
ID   AAS64474 standard; cDNA; 7821 BP.
XX
AC   AAS64474;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #278.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG00287.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity
XX
PS   Claim 1; SEQ ID No 278; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 7821 BP; 1986 A; 1904 C; 1883 G; 2046 T; 2 other;
```

```
  Query Match             56.1%;  Score 351.8;  DB 23;  Length 7821;
  Best Local Similarity   80.2%;  Pred. No. 1.6e-96;
  Matches  413;  Conservative    0;  Mismatches  102;  Indels    0;  Gaps    0;

Qy     36 tcggtttctctgctcacctccggataaatcacggggtctcccgcgccgctcatggcgcct 95
          || | ||    | |   || |||    ||   ||  ||| |  ||     |  || ||||
Db    431 tctggtttgtagatgtgcttcgggaggatgggggtctcttctccgagaagccgggagcct 490

Qy     96 cccgtccgtctcgagcgtcccttccttcccggcgctttcctgggttgcttctggcggcc 155
             |  |  |  | ||      | || | | |  ||          ||||   ||  |
Db    491 gttgggcagccggcgcccggtctcccctctgctgcggaggatccctgctggcggtcgtg 550

Qy    156 ctggtgttgctgctgtcctccttctccgatcaatgcaatgtcccggaatggcttccattt 215
           || || |   |||| |      ||  | ||||||||||| ||| ||||||||||||||
Db    551 gtgctgctcgcgctgccggtggcctggggtcaatgcaatgccccagaatggcttccattt 610

Qy    216 gccaggcctaccaacctaactgatgactttgagtttcccattgggacatatctgaactat 275
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    611 gccaggcctaccaacctaactgatgagtttgagtttcccattgggacatatctgaactat 670

Qy    276 gaatgccgccctggttattccggaagaccgttttctatcatctgcctaaaaaactcagtc 335
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    671 gaatgccgccctggttattccggaagaccgttttctatcatctgcctaaaaaactcagtc 730

Qy    336 tggacaagtgctaaggacaagtgcaaacgtaaatcatgtcgtaatcctccagatcctgtg 395
          |||||  |||||||||||| ||||| ||||||||||||||||||||||||||||||||
```

```
Db    731 tggactggtgctaaggacaggtgcagacgtaaatcatgtcgtaatcctccagatcctgtg 790

Qy    396 aatggcatggcacatgtgatcaaagacatccagttcggatcccaaattaaatattcttgt 455
          ||||||||||  |||||||||||||| ||||||||||||||||||||||||||||||||
Db    791 aatggcatggtgcatgtgatcaaaggcatccagttcggatcccaaattaaatattcttgt 850

Qy    456 cctaaaggataccgactcattggttcctcgtctgccacatgcatcatctcaggcaacact 515
          ||||||||||||||||||||||||||||||||||||||||||||||||||||  | |||
Db    851 actaaaggataccgactcattggttcctcgtctgccacatgcatcatctcaggtgatact 910

Qy    516 gtcatttgggataataaaacacctgtttgtgacag 550
          ||||||||||||||| |||||||| ||||||||||
Db    911 gtcatttgggataatgaaacacctatttgtgacag 945


RESULT   7
AAQ41867
ID   AAQ41867 standard; DNA; 6951 BP.
XX
     AAQ41867;

     14-SEP-1993  (first entry)
XX
DE   CR1 coding region.
XX
KW   C3b/C4b receptor; CR1; erythrocyte; monocyte; macrophage; granulocyte;
KW   iC3b; T cell; splenic follicular dendritic cell; soluble; complement;
KW   glomerular podocyte; B cell; C3b; C4b; inactivated C3b; phagocytotsis;
KW   plasma; ligand binding activity; immune complex; activator; allotype;
KW   endocytosis; lymphocyte; classical; alternative; pathway; cofactor; F;
KW   C3/C5 convertase; liver; cleavage; factor I; regulation; glycoprotein;
KW   S; A; B; glycosylation; duplication; repetitive intervening sequence;
KW   endoglycosidase F; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             28..6147
FT                   /*tag=  a
FT   sig_peptide     28..150
FT                   /*tag=  b
FT   mat_peptide     151..6144
FT                   /*tag=  c
FT   misc_RNA        1534..6147
FT                   /*tag=  d
FT                   /note= "CR1 cDNA"
FT   misc_difference 509..512
FT                   /*tag=  e
FT                   /note= "Unclear in the specification"
XX
PN   US5212071-A.
XX
PD   18-MAY-1993.
XX
PF   01-APR-1988;   88US-0176532.
XX
PR   01-APR-1988;   88US-0176532.
PR   03-APR-1989;   89US-0332865.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
PA   (TCEL-) T CELL SCI INC.
PA   (UYJO ) UNIV JOHNS HOPKINS.
XX
PI   Carson GR,  Concino MF,  Fearon DT,  Ip SH,  Klickstein LB;
PI   Makrides SC,  Wong WW;
XX
DR   WPI; 1993-175454/21.
DR   P-PSDB; AAR36743.
XX
PT   Nucleic acid encoding polypeptide having complement regulatory
PT   activity - used to prevent reperfusion injury, inhibit Arthus
PT   reaction and neutrophil mediated tissue damage and reduce
PT   myocardial infarct size and inflammation
XX
PS   Claim 1; Fig 1; 90pp; English.
XX
CC   This sequence represents the entire coding region for the C3b/C4b
CC   receptor (CR1).  CR1 is present on erythrocytes, monocytes/macro-
CC   phages, granulocytes, B cells, some T cells, splenic follicular
CC   dendritic cells and glomerular podocytes.  CR1 specifically binds
CC   C3b, C4b and inactivated C3b (iC3b).  A soluble form of the receptor
CC   is found in plasma which has ligand binding activity and the same
CC   molecular weight as membrane-associated CR1.  CR1 binds C3b and C4b
CC   that have covalently attached to immune complexes and other complement
CC   activators.  The consequences of these interactions depends on the
CC   type of bearing the receptor.  Erythrocyte CR1 binds immune complexes
CC   for transport to the liver.  CR1 on neutrophils and monocytes
CC   internalises bound complexes, either by adsorptive endocytosis
CC   or by phagocytotsis.  The function of CR1 on B lymphocytes is less
CC   well defined.  CR1 can inhibit the classical and alternative pathway
CC   C3/C5 convertases and act as a cofactor for the cleavage of C3b and
CC   C4b by factor I, therefore CR1 has a complement regulatory function
CC   as well as acting as a receptor.  CR1 is a glycoprotein composed of
CC   a single polypeptide chain.  Four allotypic forms of CR1 have been
CC   identified, differing by increments of approx. 40-50 kD.  The two most
CC   common forms, the F and S allotypes, also termed A and B allotypes,
CC   have molecular weights of 250 and 290 kD respectively.  The two rarer
CC   forms have molecular weights of 210 and 290 kD.  These differences
CC   represent differences in the polypeptide chain of CR1, rather than
CC   glycosylation state because they are not abolished by treatment of
CC   purified receptor portein with endoglycosidase F.  The CR1 gene has
CC   been shown to have repetitive intervening sequences which may have
CC   been duplicated in the formation of the larger allotypes.
XX
SQ   Sequence 6951 BP; 1799 A; 1692 C; 1648 G; 1807 T; 5 other;


  Query Match             55.5%;  Score 348;  DB 14;  Length 6951;
  Best Local Similarity   83.2%;  Pred. No. 2.2e-95;
  Matches  406;  Conservative    0;  Mismatches   79;  Indels    3;  Gaps    1;

Qy      66 acggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtccctttccttcc 125
           | |||| |  |  | ||    |   ||   || ||| | |    |   |||  ||   ||
Db      28 atgggggcctcttctccaagaagcccggagcctgtcgggccgccggcgcccggtctcccc 87

Qy     126 cggcgctttcctgggttgcttctggcggccctggtgttgct---gctgtcctccttctcc 182
               |||     || |  || |||||||   |||||  ||||   ||||  |      ||
Db      88 ttctgctgcggaggatccctgctggcggttgtggtgctgcttgcgctgccggtggcctgg 147

Qy     183 gatcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgac 242
           | |||||||||||  ||| ||||||||||||||||||||||||||||||||||||||||
Db     148 ggtcaatgcaatgccccagaatggcttccatttgccaggcctaccaacctaactgatgag 207

Qy     243 tttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaaga 302
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208 tttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaaga 267

Qy     303 ccgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaa 362
           |||||||||||||||||||||||||||||||||||||  ||||||||||| ||||| |
Db     268 ccgttttctatcatctgcctaaaaaactcagtctggactggtgctaaggacaggtgcaga 327

Qy     363 cgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagac 422
           |||||||||||||||||||||||||||||||||||||||||||  |||||||||||| |
Db     328 cgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggtgcatgtgatcaaaggc 387

Qy     423 atccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcc 482
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     388 atccagttcggatcccaaattaaatattcttgtactaaaggataccgactcattggttcc 447

Qy     483 tcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtt 542
           |||||||||||||||||||||||||||  | |||||||||||||||||| ||||||||||
Db     448 tcgtctgccacatgcatcatctcaggtgatactgtcatttgggataatgaaacacctatt 507

Qy     543 tgtgacag 550
           |    |||
Db     508 tnnnncag 515
```

RESULT 8
ABA09026
ID ABA09026 standard; cDNA; 7028 BP.
XX
AC ABA09026;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human CR1 protein homologue-encoding cDNA, SEQ ID NO:802.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnerary; antiulcer; ss.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03800.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457740/49.
DR P-PSDB; ABB11782.
XX
PT Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX
PS Claim 1; Page 707-709; 1963pp; English.
XX
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities; and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention.
XX
SQ Sequence 7028 BP; 1819 A; 1704 C; 1681 G; 1824 T; 0 other;

Query Match 54.4%; Score 341; DB 22; Length 7028;
Best Local Similarity 83.8%; Pred. No. 3e-93;
Matches 410; Conservative 0; Mismatches 75; Indels 4; Gaps 2;

```
Qy      66 acggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtccctttccttcc 125
           | |||| |  |  | ||    |   ||   || ||| | |    |  |||  ||  ||
Db      28 atgggggcctcttctccaagaagcccggagcctgtcgggccgccggcgccccggtctcccc 87

Qy     126 cggcgctttcctgggttgcttctggcggccctggtgttgct---gctgtcctccttctcc 182
             |||     || |  || |||||||   ||||| ||||   |||| |      ||
Db      88 ttctgctgcggaggatccctgctggcggttgtggtgctgcttgcgctgccggtggcctgg 147

Qy     183 gatcaatgcaatgtcccggaat-ggcttccatttgccaggcctaccaacctaactgatga 241
           | ||||||||||| ||| |||| ||||||||||||||||||||||||||||||||||||
Db     148 ggtcaatgcaatgccccagaatgggcttccatttgccaggcctaccaacctaactgatga 207

Qy     242 ctttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaag 301
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208 gtttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaag 267

Qy     302 accgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaa 361
           |||||||||||||||||||||||||||||||||||||  ||||||||||| |||||
Db     268 accgttttctatcatctgcctaaaaaactcagtctggactggtgctaaggacaggtgcag 327

Qy     362 acgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaaga 421
           |||||||||||||||||||||||||||||||||||||  ||||||||||||
Db     328 acgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggtgcatgtgatcaaagg 387

Qy     422 catccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttc 481
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     388 catccagttcggatcccaaattaaatattcttgtactaaaggataccgactcattggttc 447

Qy     482 ctcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgt 541
           |||||||||||||||||||||||||||  | ||||||||||||||||| ||||||| |
Db     448 ctcgtctgccacatgcatcatctcaggtgatactgtcatttgggataatgaaacacctat 507

Qy     542 ttgtgacag 550
           |||||||||
Db     508 ttgtgacag 516
```

RESULT 9
AAI60166
ID AAI60166 standard; cDNA; 7028 BP.
XX
AC AAI60166;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 4155.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

```
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM41010.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 4155; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 7028 BP; 1819 A; 1704 C; 1681 G; 1824 T; 0 other;

  Query Match             54.4%;  Score 341;  DB 22;  Length 7028;
  Best Local Similarity   83.8%;  Pred. No. 3e-93;
  Matches  410;  Conservative    0;  Mismatches   75;  Indels    4;  Gaps    2;

Qy      66 acggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtccctttccttcc 125
           | ||||  |  |  | ||    |   ||  || ||| | |    |   |||  ||   ||
Db      28 atgggggcctcttctccaagaagcccggagcctgtcgggccgccggcgcccggtctcccc 87

Qy     126 cggcgctttcctgggttgcttctggcggccctggtgttgct---gctgtcctccttctcc 182
           |||     || |  || |||||||   ||||| ||||   |||| |      ||
Db      88 ttctgctgcggaggatccctgctggcggttgtggtgctgcttgcgctgccggtggcctgg 147

Qy     183 gatcaatgcaatgtcccggaat-ggcttccatttgccaggcctaccaacctaactgatga 241
           | ||||||||||| ||| |||| ||||||||||||||||||||||||||||||||||||
Db     148 ggtcaatgcaatgccccagaatgggcttccatttgccaggcctaccaacctaactgatga 207

Qy     242 ctttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaag 301
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208 gtttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaag 267

Qy     302 accgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaa 361
           |||||||||||||||||||||||||||||||||||||||||  ||||||||||| |||||
Db     268 accgttttctatcatctgcctaaaaaactcagtctggactggtgctaaggacaggtgcag 327

Qy     362 acgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaaga 421
           |||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db     328 acgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggtgcatgtgatcaaagg 387

Qy     422 catccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttc 481
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     388 catccagttcggatcccaaattaaatattcttgtactaaaggataccgactcattggttc 447

Qy     482 ctcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgt 541
           |||||||||||||||||||||||||||  | ||||||||||||||||| ||||||||  |
Db     448 ctcgtctgccacatgcatcatctcaggtgatactgtcatttgggataatgaaacacctat 507

Qy     542 ttgtgacag 550
           |||||||||
Db     508 ttgtgacag 516


RESULT  10
AAS64290
ID   AAS64290 standard; cDNA; 9038 BP.
XX
AC   AAS64290;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #94.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG00103.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 94; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
```

```
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 9038 BP; 2312 A; 2224 C; 2095 G; 2404 T; 3 other;

  Query Match            43.6%;  Score 273.2;  DB 23;  Length 9038;
  Best Local Similarity  86.6%;  Pred. No. 1.6e-72;
  Matches  502;  Conservative    0;  Mismatches   48;  Indels   30;  Gaps   17;

Qy      1 cggactcagaagggacttccctgctcggctggctttcggtttctctgctcacctccggat 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1583 cggactcagaagggacttccctgctcggctggctttcggtttctctgctcacctccggat 1642

Qy     61 aaatcacggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtccc-ttt 119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   1643 aaatcacggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtcccttttt 1702

Qy    120 ccttcccggcgctttcct-gggttgcttctggcggccct-ggtgttgctgctgtcctcct 177
          |||||||||||||||||| |||||||||||||||||| |||||||||||||||||||||
Db   1703 ccttcccggcgctttcctgggttgcttctggcggccctgggtgttgctgctgtcctcct 1762

Qy    178 tctccgatcaat-gcaatgtcccggaatggcttccattt-----gccaggcctaccaacc 231
          |||||||||||| |||||||||||||||||  | |  ||          |     | |||
Db   1763 tctccgatcaatggcaatgtcccggaatgggcttccattttgccaggggcttaccaaacc 1822

Qy    232 taactgatgactttgagtttcccattggg--acatatctgaactatgaatgccgcccct-- 287
          |||||||||| ||||||||||||||||||        ||||||||||||||||  ||
Db   1823 taactgatgaatttgagtttcccattggggacatatcctgaactatgaatgcccgccccct 1882

Qy    288 -ggttattccggaagaccgtttt-ctatcatctg-cctaaaaaactcagtct-ggacaag 343
           | |  | |||||||||||||||| |||||||||| |||||||||||||||| |||||||
Db   1883 ggttattcccggaagaccgttttcctatcatctgccctaaaaaactcagtctgggacaag 1942

Qy    344 tgctaaggacaagtgcaaacgtaaa---tcatgtcgtaatcctccagatcctgtgaat-g 399
          |||||||||||||||||||||| ||      |     |  | ||||||||||||||| |
Db   1943 tgctaaggacaagtgcaaacgttaaatcatgttcgtaatccttccagatcctgtgaatgg 2002

Qy    400 gcatggcacatgtga--tcaaagacatccagttcggat---cccaaattaaatattcttg 454
          |||||||||||||   ||||||||||||||||  |     || || ||||||||||||
Db   2003 gcatggcacatgtggattcaaagacatccagtttcggattcccaaatttaaatattcttg 2062

      455 -tcctaaaggataccgactcattgg--ttcctcgtctgccacat-gcatcatctcaggca 510
           |||||||||||||||||||||||  |||||||||||||||| ||||||||||||||||
Db   2063 ttcctaaaggataccgactcattgggtttcctcgtctgccacatggcatcatctcaggca 2122

Qy    511 acactgtcatttgggataataaaacacctgtttgtgacag 550
          |||||||||||||||||||||||||||||||||||||||
Db   2123 acactgtcatttgggataataaaacacctgtttgtgacag 2162


RESULT  11
AAS64473
ID   AAS64473 standard; cDNA; 2796 BP.
XX
AC   AAS64473;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #277.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG00286.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 277; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2796 BP; 949 A; 571 C; 591 G; 685 T; 0 other;

  Query Match            43.0%;  Score 269.4;  DB 23;  Length 2796;
  Best Local Similarity  97.6%;  Pred. No. 1.3e-71;
  Matches  284;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy    175 ccttctccgatcaatgcaatgtcccggaat-ggcttccatttgccaggcctaccaaccta 233
          || ||    |||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1909 ccatcggaaatcaatgcaatgtcccggaatgggcttccatttgccaggcctaccaaccta 1968

Qy    234 actgatgactttgagtttcccattgggacatatctgaactatgaatgccgccctggttat 293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1969 actgatgactttgagtttcccattgggacatatctgaactatgaatgccgccctggttat 2028

Qy    294 tccggaagaccgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggac 353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2029 tccggaagaccgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggac 2088

Qy    354 aagtgcaaacgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtg 413
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2089 aagtgcaaacgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtg 2148

Qy    414 atcaaagacatccagttcggatcccaaattaaatattcttgtcctaaagga 464
          ||||||||||||||||| |||||||||||||||||||||||||||||||
Db   2149 atcaaagacatccagttcagatcccaaattaaatattcttgtcctaaagga 2199


RESULT  12
```

```
AAV53262
ID   AAV53262 standard; DNA; 591 BP.
XX
AC   AAV53262;
XX
DT   18-JAN-1999  (first entry)
XX
DE   Complement receptor type 1-like sequence CM7 DNA.
XX
KW   Complement receptor type-1; CR1; CM7; complement; inhibitor;
KW   anti-haemolytic; multiple sclerosis; Parkinson's disease;
KW   xenograft rejection; inflammation; Crohn's disease; asthma;
KW   pancreatitis; post-ischaemic reperfusion; infection; sepsis;
KW   autoimmune disease; rheumatoid arthritis; proliferative nephritis;
KW   myasthenia gravis; reproductive disorder; therapy; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9839433-A1.
XX
PD   11-SEP-1998.
XX
PF   05-MAR-1998;   98WO-GB00727.
XX
PR   05-MAR-1997;   97GB-0004519.
XX
PA   (ADPR-) ADPROTECH PLC.
XX
PI   Cox VF,  Mossakowska DEI,  Smith RAG;
XX
DR   WPI; 1998-506358/43.
DR   P-PSDB; AAW79236.
XX
PT   Soluble polypeptide comprising short consensus repeats from LHR-A -
PT   used to treat disorders and diseases associated with inflammation or
PT   inappropriate complement activation
XX
PS   Claim 22; Page 42-43; 67pp; English.
XX
CC   This DNA sequence encodes CM7 (see AAW79236), a protein that consists
CC   of the short consensus repeats (SCR) 1 and 2 from the complement
CC   receptor type 1 (CR1) fused to the SCR3 of CR1-like pseudogene (see
CC   AAW79247).  CM7 DNA was constructed using plasmid pDB1013-5, which
CC   codes for SCR1-3 of CR1, by site-directed mutagenesis using 3 pairs
CC   of oligonucleotides (see AAV53263-65) that introduced 10 amino acid
CC   changes to the native SCR3 sequence corresponding to changes
CC   observed in the CR1-like pseudogene (Cr1pse).  pBrocSCR1-3CM7
CC   carrying the CM7 DNA construct was used to transform Escherichia
CC   coli BL21(DE3), and CM7 was purified from solubilised inclusion
CC   bodies.  The invention provides DNA sequences (see AAV53262 and
CC   AAV53269-79) encoding novel soluble engineered CR1 polypeptides (see
CC   AAW53236-47) such as CM7 that act as complement inhibitors with
CC   functional complement inhibitory, including anti-haemolytic,
CC   activity.  These can be used to treat a disease or disorder
CC   associated with inflammation or inappropriate complement activation,
CC   such as neurological disorders (e.g. multiple sclerosis and
CC   Parkinson's disease), disorders of inappropriate or undesirable
CC   complement activation (e.g. xenograft rejection), inflammatory
CC   disorders (e.g. Crohn's disease, asthma, and acute pancreatitis),
CC   post-ischaemic reperfusion conditions, infection or sepsis,
CC   immune complex disorders and autoimmune diseases (e.g. rheumatoid
CC   arthritis, proliferative nephritis and myasthenia gravis), and
CC   reproductive disorders.
XX
SQ   Sequence 591 BP; 132 A; 159 C; 148 G; 152 T; 0 other;


  Query Match           42.6%;  Score 267.4;  DB 19;  Length 591;
  Best Local Similarity  83.3%;  Pred. No. 2.5e-71;
  Matches  304;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy    186 caatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgacttt 245
          || |||||  |  |||||||||||  || ||  ||  | ||  ||||||||  ||||||||  |||
Db      4 cagtgcaacgctccggaatggctgccgttcgcgcgcccgaccaacctgactgatgaatttt 63

Qy    246 gagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaagaccg 305
          |||||  ||  ||  ||  ||  ||  ||||||||  ||||||||||  ||||||   |||   |  |||
Db     64 gagttcccgatcggtacctacctgaactacgaatgccgcccgggttatagcggccgcccg 123

Qy    306 ttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaacgt 365
          ||||||||||||||||  ||||||||  ||||||||  ||||||||||   |||  ||||
Db    124 ttttctatcatctgcctgaaaaactctgtctggactggtgctaaggaccgttgccgacgt 183

Qy    366 aaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagacatc 425
          |||||  ||||||||||  ||||||||  ||  ||  |||||||   ||||||||||||  ||||
Db    184 aaatcttgtcgtaatccgccagatccggttaacggcatggtgcatgtgatcaaaggcatc 243

Qy    426 cagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcctcg 485
          ||||||||  ||||||||||||||||||||||  |||||||  |||||  ||  ||||||||||
Db    244 cagttcggttcccaaattaaatattcttgtactaaaggttaccgtctgattggttcctcc 303

Qy    486 tctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtttgt 545
             || ||||||||||||||  ||  |  ||||||||||||||||||  |||||||  |||||
Db    304 agcgctacatgcatcatctctggtgatactgtcatttgggataatgaaacaccgatttgt 363

Qy    546 gacag 550
          |||  |
Db    364 gaccg 368


RESULT  13
AAV53269
ID   AAV53269 standard; DNA; 591 BP.
XX
AC   AAV53269;
XX
DT   18-JAN-1999  (first entry)
XX
DE   Complement receptor type 1-like sequence CM1 DNA.
XX
KW   Complement receptor type-1; CR1; CM1; complement; inhibitor;
KW   anti-haemolytic; multiple sclerosis; Parkinson's disease;
KW   xenograft rejection; inflammation; Crohn's disease; asthma;
KW   pancreatitis; post-ischaemic reperfusion; infection; sepsis;
KW   autoimmune disease; rheumatoid arthritis; proliferative nephritis;
KW   myasthenia gravis; reproductive disorder; therapy; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9839433-A1.
XX
PD   11-SEP-1998.
XX
PF   05-MAR-1998;   98WO-GB00727.
XX
PR   05-MAR-1997;   97GB-0004519.
XX
PA   (ADPR-) ADPROTECH PLC.
XX
PI   Cox VF,  Mossakowska DEI,  Smith RAG;
XX
DR   WPI; 1998-506358/43.
DR   P-PSDB; AAW79237.
XX
PT   Soluble polypeptide comprising short consensus repeats from LHR-A -
PT   used to treat disorders and diseases associated with inflammation or
PT   inappropriate complement activation
XX
PS   Claim 22; Page 44; 67pp; English.
XX
CC   This DNA sequence encodes CM1 (see AAW79237), a protein that consists
CC   of the short consensus repeats (SCR) 1 and 2 from complement
CC   receptor type 1 (CR1) fused to an SCR3 (see AAW79242) in which 5 amino
CC   acids were altered to those found in the SCR3 of the CR1-like
```

```
CC   pseudogene (Crlpse) putative product.  CM1 DNA was constructed by
CC   site-directed mutagenesis (see AAV53263) of plasmid pDB1013-5, which
CC   codes for SCR1-3 of CR1.  pBrocSCR1-3CM1 carrying CM1 DNA was used
CC   to transform Escherichia coli BL21(DE3), and CM1 was purified from
CC   solubilised inclusion bodies.  The invention provides DNA sequences
CC   (see AAV53262 and AAV53269-79) encoding novel soluble engineered CR1
CC   polypeptides (see AAW53236-47) such as CM1 that act as complement
CC   inhibitors with functional complement inhibitory, including
CC   anti-haemolytic, activity.  These can be used to treat a disease or
CC   disorder associated with inflammation or inappropriate complement
CC   activation, such as neurological disorders (e.g. multiple sclerosis
CC   and Parkinson's disease), disorders of inappropriate or undesirable
CC   complement activation (e.g. xenograft rejection), inflammatory
CC   disorders (e.g. Crohn's disease, asthma, and acute pancreatitis),
CC   post-ischaemic reperfusion conditions, infection or sepsis,
CC   immune complex disorders and autoimmune diseases (e.g. rheumatoid
CC   arthritis, proliferative nephritis and myasthenia gravis), and
CC   reproductive disorders.
XX
SQ   Sequence 591 BP; 127 A; 159 C; 151 G; 154 T; 0 other;

  Query Match             42.6%;  Score 267.4;  DB 19;  Length 591;
  Best Local Similarity   83.3%;  Pred. No. 2.5e-71;
  Matches  304;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy    186 caatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgacttt 245
          || ||||| |  |||||||||||| || || ||  | || |||||||| |||||||| |||
Db      4 cagtgcaacgctccggaatggctgccgttcgcgcgcccgaccaacctgactgatgaattt 63

Qy    246 gagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaagaccg 305
          ||||| || || || || || |||||||| |||||||||| ||||||  |||  | |||
Db     64 gagttcccgatcggtacctacctgaactacgaatgccgcccgggttatagcggccgcccg 123

Qy    306 ttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaacgt 365
          |||||||||||||||| |||||||| ||||||||  ||||||||||   |||  ||||
Db    124 ttttctatcatctgcctgaaaaactctgtctggactggtgctaaggaccgttgccgacgt 183

Qy    366 aaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagacatc 425
          ||||| |||||||||||| |||||||| || || |||||||  |||||||||||| ||||
Db    184 aaatcttgtcgtaatccgccagatccggttaacggcatggtgcatgtgatcaaaggcatc 243

Qy    426 cagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcctcg 485
          |||||||| |||||||||||||||||||| ||||||| ||||| || ||||||||||
Db    244 cagttcggttcccaaattaaatattcttgtactaaaggttaccgtctgattggttcctcc 303

Qy    486 tctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtttgt 545
            || |||||||||||||| ||  | ||||||||||||||||||| |||||||  |||||
      304 agcgctacatgcatcatctctggtgatactgtcatttgggataatgaaacaccgatttgt 363

Qy    546 gacag 550
          ||| |
Db    364 gaccg 368


RESULT  14
AAV53270
ID   AAV53270 standard; DNA; 591 BP.
XX
AC   AAV53270;
XX
DT   18-JAN-1999  (first entry)
XX
DE   Complement receptor type 1-like sequence CM2 DNA.
XX
KW   Complement receptor type-1; CR1; CM2; complement; inhibitor;
KW   anti-haemolytic; multiple sclerosis; Parkinson's disease;
KW   xenograft rejection; inflammation; Crohn's disease; asthma;
KW   pancreatitis; post-ischaemic reperfusion; infection; sepsis;
KW   autoimmune disease; rheumatoid arthritis; proliferative nephritis;
KW   myasthenia gravis; reproductive disorder; therapy; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9839433-A1.
XX
PD   11-SEP-1998.
XX
PF   05-MAR-1998;   98WO-GB00727.
XX
PR   05-MAR-1997;   97GB-0004519.
XX
PA   (ADPR-) ADPROTECH PLC.
XX
PI   Cox VF,  Mossakowska DEI,  Smith RAG;
XX
DR   WPI; 1998-506358/43.
DR   P-PSDB; AAW79238.
XX
PT   Soluble polypeptide comprising short consensus repeats from LHR-A -
PT   used to treat disorders and diseases associated with inflammation or
PT   inappropriate complement activation
XX
PS   Claim 22; Page 45; 67pp; English.
XX
CC   This DNA sequence encodes CM2 (see AAW79238), a protein that consists
CC   of the short consensus repeats (SCR) 1 and 2 from complement
CC   receptor type 1 (CR1) fused to an SCR3 (see AAW79243) in which 4 amino
CC   acids were altered to those found in the SCR3 of the CR1-like
CC   pseudogene (Crlpse) putative product.  CM2 DNA was constructed by
CC   site-directed mutagenesis (see AAV53264) of plasmid pDB1013-5, which
CC   codes for SCR1-3 of CR1.  pBrocSCR1-3CM2 carrying CM2 DNA was used
CC   to transform Escherichia coli BL21(DE3), and CM2 was purified from
CC   solubilised inclusion bodies.  The invention provides DNA sequences
CC   (see AAV53262 and AAV53269-79) encoding novel soluble engineered CR1
CC   polypeptides (see AAW53236-47) such as CM2 that act as complement
CC   inhibitors with functional complement inhibitory, including
CC   anti-haemolytic, activity.  These can be used to treat a disease or
CC   disorder associated with inflammation or inappropriate complement
CC   activation, such as neurological disorders (e.g. multiple sclerosis
CC   and Parkinson's disease), disorders of inappropriate or undesirable
CC   complement activation (e.g. xenograft rejection), inflammatory
CC   disorders (e.g. Crohn's disease, asthma, and acute pancreatitis),
CC   post-ischaemic reperfusion conditions, infection or sepsis,
CC   immune complex disorders and autoimmune diseases (e.g. rheumatoid
CC   arthritis, proliferative nephritis and myasthenia gravis), and
CC   reproductive disorders.
XX
SQ   Sequence 591 BP; 134 A; 158 C; 146 G; 153 T; 0 other;

  Query Match             42.6%;  Score 267.4;  DB 19;  Length 591;
  Best Local Similarity   83.3%;  Pred. No. 2.5e-71;
  Matches  304;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy    186 caatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgacttt 245
          || ||||| |  |||||||||||| || || ||  | || |||||||| |||||||| |||
Db      4 cagtgcaacgctccggaatggctgccgttcgcgcgcccgaccaacctgactgatgaattt 63

Qy    246 gagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaagaccg 305
          ||||| || || || || || |||||||| |||||||||| ||||||  |||  | |||
Db     64 gagttcccgatcggtacctacctgaactacgaatgccgcccgggttatagcggccgcccg 123

Qy    306 ttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaacgt 365
          |||||||||||||||| |||||||| ||||||||  ||||||||||   |||  ||||
Db    124 ttttctatcatctgcctgaaaaactctgtctggactggtgctaaggaccgttgccgacgt 183

Qy    366 aaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagacatc 425
          ||||| |||||||||||| |||||||| || || |||||||  |||||||||||| ||||
Db    184 aaatcttgtcgtaatccgccagatccggttaacggcatggtgcatgtgatcaaaggcatc 243

Qy    426 cagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcctcg 485
          |||||||| |||||||||||||||||||| ||||||| ||||| || ||||||||||
Db    244 cagttcggttcccaaattaaatattcttgtactaaaggttaccgtctgattggttcctcc 303
```

```
Qy    486 tctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtttgt 545
          || ||||||||||||||| ||  | ||||||||||||||||||| |||||||  |||||
Db    304 agcgctacatgcatcatctctggtgatactgtcatttgggataatgaaacaccgatttgt 363

Qy    546 gacag 550
          ||| |
Db    364 gaccg 368


RESULT 15
AAV53271
ID   AAV53271 standard; DNA; 591 BP.
XX
AC   AAV53271;
XX
DT   18-JAN-1999  (first entry)
XX
DE   Complement receptor type 1-like sequence CM3 DNA.
XX
     Complement receptor type-1; CR1; CM3; complement; inhibitor;
     anti-haemolytic; multiple sclerosis; Parkinson's disease;
     xenograft rejection; inflammation; Crohn's disease; asthma;
KW   pancreatitis; post-ischaemic reperfusion; infection; sepsis;
KW   autoimmune disease; rheumatoid arthritis; proliferative nephritis;
KW   myasthenia gravis; reproductive disorder; therapy; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9839433-A1.
XX
PD   11-SEP-1998.
XX
PF   05-MAR-1998;   98WO-GB00727.
XX
PR   05-MAR-1997;   97GB-0004519.
XX
PA   (ADPR-) ADPROTECH PLC.
XX
PI   Cox VF,  Mossakowska DEI,  Smith RAG;
XX
DR   WPI; 1998-506358/43.
DR   P-PSDB; AAW79239.
XX
PT   Soluble polypeptide comprising short consensus repeats from LHR-A -
PT   used to treat disorders and diseases associated with inflammation or
PT   inappropriate complement activation
XX
     Claim 22; Page 46; 67pp; English.

CC   This DNA sequence encodes CM2 (see AAW79239), a protein that consists
CC   of the short consensus repeats (SCR) 1 and 2 from complement
CC   receptor type 1 (CR1) fused to an SCR3 (see AAW79244) in which 1 amino
CC   acid was altered to that found in the SCR3 of the CR1-like
CC   pseudogene (Cr1pse) putative product.  CM3 DNA was constructed by
CC   site-directed mutagenesis (see AAV53265) of plasmid pDB1013-5, which
CC   codes for SCR1-3 of CR1.  pBrocSCR1-3CM3 carrying CM3 DNA was used
CC   to transform Escherichia coli BL21(DE3), and CM3 was purified from
CC   solubilised inclusion bodies.  The invention provides DNA sequences
CC   (see AAV53262 and AAV53269-79) encoding novel soluble engineered CR1
CC   polypeptides (see AAW53236-47) such as CM3 that act as complement
CC   inhibitors with functional complement inhibitory, including
CC   anti-haemolytic, activity.  These can be used to treat a disease or
CC   disorder associated with inflammation or inappropriate complement
CC   activation, such as neurological disorders (e.g. multiple sclerosis
CC   and Parkinson's disease), disorders of inappropriate or undesirable
CC   complement activation (e.g. xenograft rejection), inflammatory
CC   disorders (e.g. Crohn's disease, asthma, and acute pancreatitis),
CC   post-ischaemic reperfusion conditions, infection or sepsis,
CC   immune complex disorders and autoimmune diseases (e.g. rheumatoid
CC   arthritis, proliferative nephritis and myasthenia gravis), and
CC   reproductive disorders.
XX

SQ   Sequence 591 BP; 131 A; 160 C; 149 G; 151 T; 0 other;


  Query Match             42.6%;  Score 267.4;  DB 19;  Length 591;
  Best Local Similarity   83.3%;  Pred. No. 2.5e-71;
  Matches  304;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy    186 caatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgacttt 245
          || ||||| |  |||||||||||| || || ||  | || |||||||| |||||||| |||
Db      4 cagtgcaacgctccggaatggctgccgttcgcgcgcccgaccaacctgactgatgaattt 63

Qy    246 gagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaagaccg 305
          ||||| || || || || || |||||||| |||||||||| ||||||  |||  | |||
Db     64 gagttcccgatcggtacctacctgaactacgaatgccgcccgggttatagcggccgcccg 123

Qy    306 ttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaacgt 365
          |||||||||||||||| ||||||| |||||||  ||||||||||   |||  ||||
Db    124 ttttctatcatctgcctgaaaaactctgtctggactggtgctaaggaccgttgccgacgt 183

Qy    366 aaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagacatc 425
          ||||| |||||||||| |||||||| || || |||||||  |||||||||||| ||||
Db    184 aaatcttgtcgtaatccgccagatccggttaacggcatggtgcatgtgatcaaaggcatc 243

Qy    426 cagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcctcg 485
          |||||||| |||||||||||||||||||| ||||||| ||||| || ||||||||||
Db    244 cagttcggttcccaaattaaatattcttgtactaaaggttaccgtctgattggttcctcc 303

Qy    486 tctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtttgt 545
          || ||||||||||||||| ||  | ||||||||||||||||||| |||||||  |||||
Db    304 agcgctacatgcatcatctctggtgatactgtcatttgggataatgaaacaccgatttgt 363

Qy    546 gacag 550
          ||| |
Db    364 gaccg 368


Search completed: October  9, 2002, 17:40:21
Job time: 3772 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 9, 2002, 16:04:33 ; Search time 1594.28 Seconds
(without alignments)
5308.096 Million cell updates/sec

Title: US-10-031-904-30
Perfect score: 627
Sequence: 1 cggactcagaagggacttcc.....ataataaaaatcttaaccga 627

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	445	71.0	465	10	BE552138	BE552138 hw29d02.x
2	443.6	70.7	832	10	BF240184	BF240184 601905704
c 3	435	69.4	443	9	AI240881	AI240881 qj94e04.x
c 4	393.6	62.8	444	9	AI718588	AI718588 as46h01.x
5	352	56.1	1063	10	BM477528	BM477528 AGENCOURT
c 6	344.8	55.0	356	9	AI735459	AI735459 at11h03.x
7	326.8	52.1	440	10	H73873	H73873 ys14d08.r1
8	307	49.0	346	10	T66824	T66824 ya50b08.r3
c 9	306.4	48.9	389	10	T66823	T66823 ya50b08.s3
10	270.4	43.1	541	10	T83269	T83269 yd41a11.r1
c 11	268.6	42.8	433	9	AI735085	AI735085 as44f09.x
12	212.8	33.9	577	10	BG077250	BG077250 H3013A12-
13	212.8	33.9	614	9	AA107525	AA107525 mp08b07.r
14	212.8	33.9	658	10	BI697900	BI697900 603344033
15	212.8	33.9	795	10	BI455761	BI455761 603171769
16	212.8	33.9	807	10	BI078162	BI078162 602872480
17	211.2	33.7	601	9	AA212152	AA212152 mu80g07.r
18	208	33.2	773	10	BI556660	BI556660 603239267
19	206.8	33.0	722	10	BI696340	BI696340 603345540
20	205	32.7	488	10	BF023537	BF023537 ux09c08.y
21	205	32.7	569	10	BI693880	BI693880 603342124
22	205	32.7	684	9	BB613937	BB613937 BB613937
23	204.8	32.7	643	9	AW912091	AW912091 uf41f04.y
24	203.4	32.4	525	9	AA286570	AA286570 vb76a03.r
25	203.4	32.4	650	9	AW610808	AW610808 up37f12.x
26	201.8	32.2	1646	11	AK004825	AK004825 Mus muscu
27	199.6	31.8	355	10	T27695	T27695 EST12512 Hu
28	195	31.1	739	10	BG243080	BG243080 602355374
29	194.6	31.0	861	10	BF178335	BF178335 601808574
30	194	30.9	388	10	BG146321	BG146321 mab90c07.
31	194	30.9	403	9	AW610947	AW610947 un74b11.y
32	194	30.9	409	9	AW012541	AW012541 uo06c01.y
33	192.8	30.7	441	9	AW825298	AW825298 us13d06.y
34	189.2	30.2	303	10	R01561	R01561 ye79h05.r1
35	184.6	29.4	469	9	AA261275	AA261275 vb03b10.r
36	184.6	29.4	981	10	BE917021	BE917021 601663878
37	184.2	29.4	889	10	BF166939	BF166939 601774087
38	182.6	29.1	387	9	AW211552	AW211552 uo82e02.y
39	180	28.7	519	10	BF159845	BF159845 601767506
40	170.2	27.1	431	10	BF015454	BF015454 uy21g02.y
41	166	26.5	770	10	BF124847	BF124847 601756853
42	164.4	26.2	3859	11	AK017702	AK017702 Mus muscu
43	159.4	25.4	419	9	AA756782	AA756782 vv65f04.r
44	157	25.0	397	9	AA314773	AA314773 EST186595
45	156.4	24.9	300	9	AU100130	AU100130 AU100130

ALIGNMENTS

RESULT 1
BE552138/c
LOCUS BE552138 465 bp mRNA linear EST 10-AUG-2000
DEFINITION hw29d02.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3184323 3' similar to gb:Y00816_cds1 COMPLEMENT RECEPTOR TYPE 1 PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION BE552138
VERSION BE552138.1 GI:9793830
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 465)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 456.
FEATURES Location/Qualifiers
source 1. .465
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3184323"
/clone_lib="NCI_CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid3 was

prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 134 a 85 c 107 g 139 t
ORIGIN

Query Match 71.0%; Score 445; DB 10; Length 465;
Best Local Similarity 100.0%; Pred. No. 9.5e-113;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    183 gatcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgac 242
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    453 GATCAATGCAATGTCCCGGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAC 394

Qy    243 tttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaaga 302
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    393 TTTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGA 334

Qy    303 ccgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaa 362
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    333 CCGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACAAGTGCTAAGGACAAGTGCAAA 274

Qy    363 cgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagac 422
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    273 CGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATGGCACATGTGATCAAAGAC 214

Qy    423 atccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcc 482
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    213 ATCCAGTTCGGATCCCAAATTAAATATTCTTGTCCTAAAGGATACCGACTCATTGGTTCC 154

Qy    483 tcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtt 542
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    153 TCGTCTGCCACATGCATCATCTCAGGCAACACTGTCATTTGGGATAATAAAACACCTGTT 94

Qy    543 tgtgacagtgagttgaaatatgcattcctatttcttttaccgatacattctaatttttct 602
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     93 TGTGACAGTGAGTTGAAATATGCATTCCTATTTCTTTTACCGATACATTCTAATTTTTCT 34

Qy    603 ctggaataataaaaatcttaaccga 627
          |||||||||||||||||||||||||
Db     33 CTGGAATAATAAAAATCTTAACCGA 9
```

RESULT 2
AI240184
LOCUS BF240184 832 bp mRNA linear EST 14-NOV-2000
DEFINITION 601905704F1 NIH_MGC_54 Homo sapiens cDNA clone IMAGE:4133424 5', mRNA sequence.
ACCESSION BF240184
VERSION BF240184.1 GI:11154107
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 832)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1034 row: h column: 01
High quality sequence stop: 544.

FEATURES Location/Qualifiers
source 1. .832
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4133424"
/clone_lib="NIH_MGC_54"
/tissue_type="from chronic myelogenous leukemia"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: bone marrow; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 188 a 226 c 197 g 220 t 1 others
ORIGIN

Query Match 70.7%; Score 443.6; DB 10; Length 832;
Best Local Similarity 95.2%; Pred. No. 2.9e-112;
Matches 521; Conservative 0; Mismatches 20; Indels 6; Gaps 6;

```
Qy      1 cggactcagaagggacttccctgctcggctggctttcggtttctctgctcacctccggat 60
          |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db     29 CGGACTCAGAAGGGACTTCCCTGCTCGG-TGGCTTTCGGTTTCTCTGCTCACCTCCGGAT 87

Qy     61 aaatcacggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtccctttc 120
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     88 AAATCACGGGGTCTCCCGCGCCG-TCATGGCGCCTCCCGTCCGTCTCGAGCGTCCCTTTC 146

Qy    121 cttcccggcgctttcctgggttgcttctggcggccctggtgttgctgctgtcctccttct 180
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    147 CTTCCCGGCGCTTTCCTGGGTTGCTTCTGGCGG-CCTGGTGTTGCTGCTGTCCTCCTTCT 205

Qy    181 ccgatcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatg 240
          |||||||||||||||||||||||||||||||||||||||||   | ||||||||||||||
Db    206 CCGATCAATGCAATGTCCCGGAATGGCTTCCATTTGCCAGGCTACCAAACCTAACTGATG 265

Qy    241 actttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaa 300
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    266 ACTTTGAGTTTCCCATTNGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAA 325

Qy    301 gaccgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgca 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    326 GACCGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACAAGTGCTAAGGACAAGTGCA 385

Qy    361 aacgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggca-catgtgatcaaa 419
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    386 AACGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATGGCACCATGTGATCAAA 445

Qy    420 gacatccagttc-ggatcccaaattaaatattcttgtcctaaaggataccgactcattgg 478
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    446 GACATCCAGTTCAAGATCCCAAATTAAATATTCTTGTCCTAAAGGATACCGACTCATTGG 505

Qy    479 ttcctcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacc 538
          |||||||||||| |||||||||||||||||||||||||||||||  ||  |   | | | |
Db    506 TTCCTCGTCTG-CACATGCATCATCTCAGGCAACACTGTCATTCTGGGCATACATAAAAC 564

Qy    539 tgtttgt 545
          |||
Db    565 CACCTGT 571
```

RESULT 3
AI240881/c
LOCUS AI240881 443 bp mRNA linear EST 28-JAN-1999

```
DEFINITION  qj94e04.x1 NCI_CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1867134 3'
            similar to gb:Y00816_cds1 COMPLEMENT RECEPTOR TYPE 1 PRECURSOR
            (HUMAN);, mRNA sequence.
ACCESSION   AI240881
VERSION     AI240881.1  GI:3836278
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 443)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 625   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 373.
FEATURES             Location/Qualifiers
     source          1. .443
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1867134"
                     /clone_lib="NCI_CGAP_Kid3"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer,
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia),  digested with Not I and cloned into the Not
                     I and Eco RI sites  of the modified pT7T3 vector.  mRNA
                     source: 2 pooled kidneys. Library went through one round
                     of normalization. Library constructed by Bento Soares and
                     M. Fatima Bonaldo.  "
BASE COUNT      131 a     80 c    104 g    128 t
ORIGIN

  Query Match             69.4%;  Score 435;  DB 9;  Length 443;
  Best Local Similarity   98.9%;  Pred. No. 5.6e-110;
  Matches  438;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy    185 tcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgacttt 244
          ||||||||||||||| | |  ||||||||||||||||||||||||||||||||||||||||
Db    443 TCAATGCAATGTCCGGAATGGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGACTC 384

Qy    245 tgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaagacc 304
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    383 TGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGACC 324

Qy    305 gttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaacg 364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    323 GTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACAAGTGCTAAGGACAAGTGCAAACG 264

Qy    365 taaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagacat 424
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263 TAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATGGCACATGTGATCAAAGACAT 204

Qy    425 ccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcctc 484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 CCAGTTCGGATCCCAAATTAAATATTCTTGTCCTAAAGGATACCGACTCATTGGTTCCTC 144

Qy    485 gtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtttg 544
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    143 GTCTGCCACATGCATCATCTCAGGCAACACTGTCATTTGGGATAATAAAACACCTGTTTG 84

Qy    545 tgacagtgagttgaaatatgcattcctatttcttttaccgatacattctaatttttctct 604
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     83 TGACAGTGAGTTGAAATATGCATTCCTATTTCTTTTACCGATACATTCTAATTTTTCTCT 24

Qy    605 ggaataataaaaatcttaaccga 627
          |||||||||||||||||||||||
Db     23 GGAATAATAAAAATCTTAACCGA 1


RESULT   4
AI718588/c
LOCUS       AI718588                444 bp    mRNA    linear   EST 10-JUN-1999
DEFINITION  as46h01.x1 Barstead aorta HPLRB6 Homo sapiens cDNA clone
            IMAGE:2320273 3' similar to gb:Y00816_cds1 COMPLEMENT RECEPTOR TYPE
            1 PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   AI718588
VERSION     AI718588.1  GI:5035844
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 444)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .444
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2320273"
                     /clone_lib="Barstead aorta HPLRB6"
                     /sex="male"
                     /dev_stage="adult, age 64"
                     /lab_host="DH10B (phage resistant)"
                     /note="Organ: aorta; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [5' AATTCGGATCGAAC 3' and 5' GTTGGATCGG 3'], digested
                     with Not I and cloned into the Not I and Eco RI sites of
                     the modified pT7T3 vector. Library constructed by Bob
                     Barstead."
BASE COUNT      134 a     87 c    103 g    120 t
ORIGIN

  Query Match             62.8%;  Score 393.6;  DB 9;  Length 444;
  Best Local Similarity   94.4%;  Pred. No. 1.7e-98;
  Matches  419;  Conservative    0;  Mismatches   24;  Indels    1;  Gaps    1;

Qy    183 gatcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgac 242
          | |||||||||||| ||| ||||||||||||||||||||||||||||||||||||||||||
Db    444 GGTCAATGCAATGCCCCAGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAG 385

Qy    243 tttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaaga 302
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    384 TTTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGA 325

Qy    303 ccgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaa 362
          ||||||||||||||||||||||||||||||||||||||||||  ||||||||||| ||||| |
Db    324 CCGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACTGGTGCTAAGGACAGGTGCAGA 265

Qy    363 cgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagac 422
          ||||||||||||||||||||||||||||||||||||||||||  |||||||||||| | |
Db    264 CGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATGGTGCATGTGATCAAAGGC 205

Qy    423 atccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcc 482
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    204 ATCCAGTTCGGATCCCAAATTAAATATTCTTGTACTAAAGGATACCGACTCATTGGTTCC 145

Qy    483 tcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtt 542
          ||||| |||||||||||||||||||  | ||||||||||||||||| |||||||| ||
Db    144 TCGTC-GCCACATGCATCATCTCAGGTGATACTGTCATTTGGGATAATGAAACACCTATT 86

Qy    543 tgtgacagtgagttgaaatatgcattcctatttcttttaccgatacattctaatttttct 602
          |||||||||||||||||||| | ||||||||||||||||||  ||||||||||||||||
Db     85 TGTGACAGTGAGTTGAAATATCCCTTCCTATTTCTTTTACCGACCCATTCTAATTTTTCT 26

Qy    603 ctggaataataaaaatcttaaccg 626
          |||||||||||||||||  || |
Db     25 CTGGAATAATAAAAATCTATACTG 2


RESULT   5
BM477528
LOCUS       BM477528                1063 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6484929 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5554371
            5', mRNA sequence.
ACCESSION   BM477528
VERSION     BM477528.1  GI:18526570
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1063)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12272  row: n  column: 04
            High quality sequence stop: 749.
FEATURES             Location/Qualifiers
     source          1. .1063
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5554371"
                     /clone_lib="NIH_MGC_85"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      251 a    286 c    272 g    249 t      5 others
ORIGIN


  Query Match             56.1%;  Score 352;  DB 10;  Length 1063;
  Best Local Similarity   84.0%;  Pred. No. 8.3e-87;
  Matches  410;  Conservative    0;  Mismatches   75;  Indels    3;  Gaps    1;

Qy     66 acggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtccctttccttcc 125
          |||| |  |  | ||    |   ||   || ||| | |    |   |||  ||   ||
Db     19 ATGGGGGCCTCTTCTCCAAGAAGCCCGGAGCCTGTCGGGCCGCCGGCGCCCGGTCTCCCC 78

Qy    126 cggcgctttcctgggttgcttctggcggccctggtgttgct---gctgtcctccttctcc 182
             |||      || |  || |||||||   ||||| ||||   |||| |      ||
Db     79 TTCTGCTGCGGAGGATCCCTGCTGGCGGTTGTGGTGCTGCTTGCGCTGCCGGTGGCCTGG 138

Qy    183 gatcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgac 242
          | |||||||||| ||| ||||||||||||||||||||||||||||||||||||||||||
Db    139 GGTCAATGCAATGCCCCAGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAG 198

Qy    243 tttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaaga 302
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    199 TTTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGA 258

Qy    303 ccgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaa 362
          ||||||||||||||||||||||||||||||||||  |||||||||||| ||||| |
Db    259 CCGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACTGGTGCTAAGGACAGGTGCAGA 318

Qy    363 cgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagac 422
          ||||||||||||||||||||||||||||||||||||||||||  |||||||||||| | |
Db    319 CGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATGGTGCATGTGATCAAAGGC 378

Qy    423 atccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcc 482
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    379 ATCCAGTTCGGATCCCAAATTAAATATTCTTGTACTAAAGGATACCGACTCATTGGTTCC 438

Qy    483 tcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtt 542
          |||||||||||||||||||| | ||||||||||||||||| |||||||| ||
Db    439 TCGTCTGCCACATGCATCATCTCAGGTGATACTGTCATTTGGGATAATGAAACACCTATT 498

Qy    543 tgtgacag 550
          ||||||||
Db    499 TGTGACAG 506


RESULT   6
AI735459/c
LOCUS       AI735459                 356 bp    mRNA    linear   EST 14-JUN-1999
DEFINITION  at11h03.x1 Barstead aorta HPLRB6 Homo sapiens cDNA clone
            IMAGE:2354837 3' similar to gb:Y00816_cds1 COMPLEMENT RECEPTOR TYPE
            1 PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   AI735459
VERSION     AI735459.1  GI:5056983
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 356)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .356
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2354837"
```

/clone_lib="Barstead aorta HPLRB6"
/sex="male"
/dev_stage="adult, age 64"
/lab_host="DH10B (phage resistant)"
/note="Organ: aorta; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [5' AATTCGGATCGAAC 3' and 5' GTTGGATCGG 3'], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library constructed by Bob Barstead."

BASE COUNT 107 a 61 c 77 g 111 t
ORIGIN

Query Match 55.0%; Score 344.8; DB 9; Length 356;
Best Local Similarity 98.0%; Pred. No. 5.6e-85;
Matches 349; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy    266 tctgaactatgaatgccgccctggttattccggaagaccgttttctatcatctgcctaaa 325
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    356 TCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGACCGTTTTCTATCATCTGCCTAAA 297

Qy    326 aaactcagtctggacaagtgctaaggacaagtgcaaacgtaaatcatgtcgtaatcctcc 385
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    296 AAACTCAGTCTGGACAAGTGCTAAGGACAAGTGCAAACGTAAATCATGTCGTAATCCTCC 237

Qy    386 agatcctgtgaatggcatggcacatgtgatcaaagacatccagttcggatcccaaattaa 445
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    236 AGATCCTGTGAATGGCATGGCACATGTGATCAAAGACATCCAGTTCAGATCCCAAATTAA 177

Qy    446 atattcttgtcctaaaggataccgactcattggttcctcgtctgccacatgcatcatctc 505
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    176 ATATTCTTGTCCTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCATCTC 117

Qy    506 aggcaacactgtcatttgggataataaaacacctgtttgtgacagtgagttgaaatatgc 565
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    116 AGGCAACACTGTCATTTGGGATAATAAAACACCTGTTTGTGACAGTGAGTTGAAATATGC 57

Qy    566 attcctatttcttttaccgatacattctaatttttctctggaataataaaaatctt 621
          |||||||||||||  |  |||||||||||||||||||||||||||||||  ||
Db     56 ATTCCTATTTCTTAAAAAGATACATTCTAATTTTTCTCTGGAATAATAAAAAATTT 1
```

RESULT 7
H73873
LOCUS H73873 440 bp mRNA linear EST 31-OCT-1995
DEFINITION ys14d08.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:214767 5' similar to gb:Y00816_cds1 COMPLEMENT RECEPTOR TYPE 1 PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION H73873
VERSION H73873.1 GI:1046874
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 440)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 706
High quality sequence stops: 275
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 706 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 275.

FEATURES Location/Qualifiers
source 1. .440
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:214767"
/clone_lib="Soares fetal liver spleen 1NFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 83 a 128 c 100 g 125 t 4 others
ORIGIN

Query Match 52.1%; Score 326.8; DB 10; Length 440;
Best Local Similarity 98.4%; Pred. No. 6e-80;
Matches 361; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

```
Qy      4 actcagaagggacttccctgctcggctggctttcggtttctctgctcacctccggataaa 63
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2 ACTCAGAAGGNACTTCCCTGCTCGGCTGGCTTTCGGTTTCTCTGCTCACCTCCGGATAAA 61

Qy     64 tcacggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtccctttcctt 123
          |||||||||||||||||||||||||||| | ||||||||||||||||||||||||||||||
Db     62 TCACGGGGTCTCCCGCGCCGCTCATGGC-CTTCCCGTCCGTCTCGAGCGTCCCTTTCCTT 120

Qy    124 cccggcgctttcctgggttgcttctggcggccctggtgttgctgctgtcctccttctccg 183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCCGGCGCTTTCCTGGGTTGCTTCTGGCGGCCCTGGTGTTGCTGCTGTCCTCCTTCTCCG 180

Qy    184 atcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgact 243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ATCAATGCAATGTCCCGGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGACT 240

Qy    244 ttgagtttcccattgggacatatctgaactatgaatgccgccct-ggttattccggaaga 302
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    241 TTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGGTTATTCCGGAAGA 300

Qy    303 ccgttttctatcatctgcctaaaaaactcagtct-ggacaagtgctaaggacaagtgcaa 361
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    301 CCGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGGACAAGTGCTAAGGACAAGTGCAA 360

Qy    362 acgtaaa 368
          |||| ||
Db    361 ACGTTAA 367
```

RESULT 8
T66824
LOCUS T66824 346 bp mRNA linear EST 07-MAR-1995
DEFINITION ya50b08.r3 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:66327 5' similar to gb:Y00816_cds1 COMPLEMENT RECEPTOR TYPE 1 PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION T66824
VERSION T66824.1 GI:676264

```
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 346)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Other_ESTs: ya50b08.s3.exp
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High qality sequence stops: 324
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: T7
            High quality sequence stop: 324.
FEATURES             Location/Qualifiers
     source          1. .346
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:66327"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       67 a    101 c     77 g     97 t      4 others
ORIGIN


  Query Match             49.0%;  Score 307;  DB 10;  Length 346;
  Best Local Similarity   98.3%;  Pred. No. 1.7e-74;
  Matches  340;  Conservative    0;  Mismatches    3;  Indels    3;  Gaps    3;

Qy      72 tctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtccctttccttcccggcgc 131
           |||||||||||||||||| | ||||||||||||||||||||||||||||||||||||||
Db       1 TCTCCCGCGCCGCTCATGGC-CNTCCCGTCCGTCTCGAGCGTCCCTTTCCTTCCCGGCGC 59

Qy     132 tttcctgggttgcttctggcggccctggtgttgctgctgtcctccttctccgatcaatgc 191
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60 TTTCCTGGGTTGCTTCTGGCGGCCCTGGTGTTGCTGCTGTCCTCCTTCTCCGATCAATGC 119

Qy     192 aatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgactttgagttt 251
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     120 AATGTCCCGGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGACTTTGAGTTT 179

Qy     252 cccattgggacatatctgaactatgaatgccgccct-ggttattccggaagaccgttttc 310
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     180 CCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGGTTATTCCGGAAGACCGTTTTC 239

Qy     311 tatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaacgt-aaat 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||| |
Db     240 TATCATCTGCCTAAAAAACTCAGTCTGGACAAGTGCTAAGGACAAGTGCAAACGTAAAAT 299

Qy     370 catgtcgtaatcctccagatcctgtgaatggcatggcacatgtgat 415
           |||||||||||||||||||||||||||||| |||| |||||||||
Db     300 CATGTCGTAATCCTCCAGATCCTGTGAATGGNATGGNACATGTGAT 345


RESULT   9
T66823/c
LOCUS       T66823                   389 bp    mRNA    linear   EST 07-MAR-1995
DEFINITION  ya50b08.s3 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:66327 3' similar to gb:Y00816_cds1 COMPLEMENT RECEPTOR TYPE 1
            PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   T66823
VERSION     T66823.1  GI:676263
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 389)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Other_ESTs: ya50b08.r3.exp
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -21M13LR
            High quality sequence stop: 390.
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:66327"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      114 a     67 c     92 g    112 t      4 others
ORIGIN


  Query Match             48.9%;  Score 306.4;  DB 10;  Length 389;
  Best Local Similarity   95.8%;  Pred. No. 2.6e-74;
  Matches  345;  Conservative    0;  Mismatches   10;  Indels    5;  Gaps    3;

Qy     267 ctgaactatgaatgc---cgccctggttattccggaagaccgtttt-ctatcatctg-cc 321
           |||||||||||||||   | ||  | |  | |||||||||||||| |||||||||| ||
Db     369 CTGAACTATGAATGCCCGCCCCTGGTTANTCCCGGAAGACCGTTTTCCTATCATCTGCCC 310

Qy     322 taaaaaactcagtctggacaagtgctaaggacaagtgcaaacgtaaatcatgtcgtaatc 381
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     309 TAAAAAACTCAGTCTGGACAAGTGCTAAGGACAAGTGCAAACGTAAATCATGTCGTAANC 250

Qy     382 ctccagatcctgtgaatggcatggcacatgtgatcaaagacatccagttcggatcccaaa 441
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     249 CTCCAGANCCTGTGAATGGCATGGCACATGTGATCAAAGACATCCAGTTCGGATCCCAAA 190
```

```
Qy    442 ttaaatattcttgtcctaaaggataccgactcattggttcctcgtctgccacatgcatca 501
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    189 TTAAATANTCTTGTCCTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCA 130

Qy    502 tctcaggcaacactgtcatttgggataataaaacacctgtttgtgacagtgagttgaaat 561
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    129 TCTCAGGCAACACTGTCATTTGGGATAATAAAACACCTGTTTGTGACAGTGAGTTGAAAT 70

Qy    562 atgcattcctatttcttttaccgatacattctaatttttctctggaataataaaaatctt 621
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     69 ATGCATTCCTATTTCTTTTACCGATACATTCTAATTTTTCTCTGGAATAATAAAAATCTT 10
```

RESULT 10
T83269
LOCUS T83269 541 bp mRNA linear EST 16-MAR-1995
DEFINITION yd41a11.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:110780 5' similar to gb:Y00816_cds1 COMPLEMENT RECEPTOR TYPE 1 PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION T83269
VERSION T83269.1 GI:711557
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 541)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 791
High quality sequence stops: 300 Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 791 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 300.

FEATURES Location/Qualifiers
source 1. .541
/organism="Homo sapiens"
/db_xref="GDB:466397"
/db_xref="taxon:9606"
/clone="IMAGE:110780"
/clone_lib="Soares fetal liver spleen 1NFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 122 a 121 c 110 g 178 t 10 others
ORIGIN

Query Match 43.1%; Score 270.4; DB 10; Length 541;
Best Local Similarity 93.9%; Pred. No. 2.9e-64;
Matches 336; Conservative 0; Mismatches 13; Indels 9; Gaps 5;

```
Qy    132 tttcctgggttgcttctggcggccctggtgttgctgctgtcctccttctccgatcaatgc 191
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5 TTTCCTGGGTTGCTTCTGGCGGCCCTGGTGTTGCTGCTGTCCTCCTTCTCCGATCAATGC 64

Qy    192 aatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgactttgagttt 251
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     65 AATGTCCCGGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGACTTTGAGTTT 124

Qy    252 cccattgggacatatctgaactatgaatgccgccctggttattccggaagaccgttttct 311
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    125 CCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGACCGTTTTCT 184

Qy    312 atcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaacgt-aaatc 370
          ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||| 
Db    185 ATCATCTGCCTAAAAAACTCAGTCTGGACAAGTGCTAAGGACAAGTGCAAACGTNAAATC 244

Qy    371 atgtcgtaatcctccagatcctgtgaat-ggcatggcacatgtg--atcaaagacatcca 427
          |||||||||||||||||||||||||||| |||||||||||||||   ||||||||||||
Db    245 ATGTCGTAATCCTCCAGATCCTGTGAATGGGCATGGCACATGTGGATTCAAAGACATCCA 304

Qy    428 gttcgg---atcccaaattaaatattcttgt--cctaaaggataccgactcattggtt 480
          |||  |     || || ||||||||||||||  |  || || |||||| ||||||| |
Db    305 GTTTCGGATTCCCAAATTTAAATATTCTTGTTCCTAAAGGGNTACCGANTCATTGGGT 362
```

RESULT 11
AI735085/c
LOCUS AI735085 433 bp mRNA linear EST 14-JUN-1999
DEFINITION as44f09.x1 Barstead aorta HPLRB6 Homo sapiens cDNA clone IMAGE:2320073 3' similar to gb:Y00816_cds1 COMPLEMENT RECEPTOR TYPE 1 PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION AI735085
VERSION AI735085.1 GI:5056545
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 433)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 60.

FEATURES Location/Qualifiers
source 1. .433
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2320073"
/clone_lib="Barstead aorta HPLRB6"
/sex="male"
/dev_stage="adult, age 64"
/lab_host="DH10B (phage resistant)"
/note="Organ: aorta; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [5' AATTCGGATCGAAC 3' and 5' GTTGGATCGG 3'], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library constructed by Bob Barstead."

```
BASE COUNT      137 a     86 c     95 g    115 t
ORIGIN

  Query Match             42.8%;  Score 268.6;  DB 9;  Length 433;
  Best Local Similarity   81.8%;  Pred. No. 8.4e-64;
  Matches  346;  Conservative    0;  Mismatches   74;  Indels    3;  Gaps    3;

Qy    197 cccggaatggcttccatttgccaggcctaccaacctaactgatgactttgagtttcccat 256
          ||| |||||  | ||||  |  ||||  ||| |||||  |   || ||||||||||||  |
Db    430 CCCCAAATGGGCTTCATTGCCAGGGCCACCCACCCTAATTATAGAGTTTGAGTTTCCCTT 371

Qy    257 tgggacatatctgaactatgaatgccgccctggttattccggaagaccgttttctatcat 316
          |  |||||||| |||  ||| ||  ||||||||||||||| |||||  |||||||   ||||
Db    370 T-GGACATATTTGACATATAAAGCCCGCCCTGGTTATTCGGGAAGCCCGTTTTTATTCAT 312

Qy    317 ctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaacgtaaatcatgtcg 376
          || || |||||  |  || |  |  |  | |  | ||||||  ||| ||  |||| ||| |
Db    311 TTGGCTTAAAAAATTCGTTTTGGCTGGGGGTTAAGACAAGGCCAACCGGTAATCCTGTTG 252

          377 taatcctccagatcctgtgaatggcatggcacatgtgatcaaagacatccagttcggatc 436
            || |||| | | |   ||||||||||  ||||||||||||  ||||||||  |||||
Db    251 GTATTCTCCCGGTTCCTGGAATGGCATGGTGCATGTGATCAAAGGCATCCAGTCGGGATC 192

Qy    437 cc-aaattaaatattcttgtcctaaaggataccgactcattggttcctcgtctgccacat 495
          || |||||||||||||||  |||||||||||||||||||||||||||||||  |||||||
Db    191 CCAAAATTAAATATTCTTGTACTAAAGGATACCGACTCATTGGTTCCTCGTC-GCCACAT 133

Qy    496 gcatcatctcaggcaacactgtcatttgggataataaaacacctgtttgtgacagtgagt 555
          ||||||||||||  | ||||||||||||||||||  |||||||| ||||||||||||||
Db    132 GCATCATCTCAGGTGATACTGTCATTTGGGATAATGAAACACCTATTTGTGACAGTGAGT 73

Qy    556 tgaaatatgcattcctatttcttttaccgatacattctaatttttctctggaataataaa 615
          |||||||| | ||||||||||||||||||| ||||||||||||||||||||||||||||
Db     72 TGAAATATCCCTTCCTATTTCTTTTACCGACACATTCTAATTTTTCTCTGGAATAATAAA 13

Qy    616 aat 618
          |||
Db     12 AAT 10


RESULT  12
BG077250
LOCUS       BG077250                577 bp    mRNA    linear   EST 26-JAN-2001
DEFINITION  H3013A12-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
            H3013A12 5', mRNA sequence.
ACCESSION   BG077250
VERSION     BG077250.1  GI:12559818
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 577)
  AUTHORS   Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka
            ,T.S., Carter,M.G. and Ko,M.S.H.
  TITLE     Verification and initial annotation of NIA mouse 15K cDNA clone set
  JOURNAL   Unpublished (2001)
COMMENT     Other_ESTs: H3013A12-3
            Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            This clone set has been freely distributed to the community. Please
            visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details.
            Plate: H3013  row: A  column: 12
            Seq primer: -21M13 Reverse
            High quality sequence stop: 577
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .577
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="niaEST:H3013A12-5"
                     /db_xref="taxon:10090"
                     /clone="H3013A12"
                     /clone_lib="NIA Mouse 15K cDNA Clone Set"
                     /sex="Clones arrayed from a variety of cDNA libraries"
                     /dev_stage="Clones arrayed from a variety of cDNA
                     libraries"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
                     clone is among a rearrayed set of 15,247 clones from 11
                     embryo  cDNA libraries (including preimplantation stage
                     embryos from  unfertilized egg to blastocyst, embryonic
                     part of E7.5 embryos,  extraembryonic part of E7.5 embryos
                     , and E12.5 female mesonephros/gonad) and one newborn
                     ovary cDNA library. Average insert size 1.5 kb. All
                     source libraries are cloned unidirectionally with Oligo(dT
                     )-Not primers.  References include: (1) Genome-wide
                     expression profiling of mid-gestation  placenta and embryo
                     using a 15,000 mouse developmental cDNA microarray,  2000,
                     Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2)
                     Large-scale cDNA  analysis reveals phased gene expression
                     patterns during preimplantation  mouse develolpment, 2000,
                     Development, 127: 1737-1749; (3) Genome-wide  mapping of
                     unselected transcripts from extraembryonic tissue of
                     7.5-day  mouse embryos reveals enrichment in the t-complex
                     and under-representation  on the X chromosome, 1998, Hum
                     Mol Genet 7: 1967-1978."
BASE COUNT      141 a    132 c    132 g    172 t
ORIGIN

  Query Match             33.9%;  Score 212.8;  DB 10;  Length 577;
  Best Local Similarity   71.4%;  Pred. No. 2.9e-48;
  Matches  280;  Conservative    0;  Mismatches  112;  Indels    0;  Gaps    0;

Qy    156 ctggtgttgctgctgtcctccttctccgatcaatgcaatgtcccggaatggcttccattt 215
          ||| ||||  ||||| | |     |  | ||| |||   | |||   |  |||||| | |
Db    143 CTGCTGTTCTTGCTGCCATTTACTTTGGGTCACTGCCCAGCCCCATCACAGCTTCCTTCT 202

Qy    216 gccaggcctaccaacctaactgatgactttgagtttcccattgggacatatctgaactat 275
          ||||  ||||  || ||||||||||| |    ||||||||||| |||||| ||   |||
Db    203 GCCAAACCTATAAATCTAACTGATGAATCCATGTTTCCCATTGGAACATATTTGTTGTAT 262

Qy    276 gaatgccgccctggttattccggaagaccgttttctatcatctgcctaaaaaactcagtc 335
          ||||| | ||| || |||  |   || | ||| |||||||| ||||  | || |||||  |
Db    263 GAATGTCTCCCAGGATATATCAAGAGGCAGTTCTCTATCACCTGCAAACAAGACTCAACC 322

Qy    336 tggacaagtgctaaggacaagtgcaaacgtaaatcatgtcgtaatcctccagatcctgtg 395
          ||||| |||||| | || ||||| | ||| |||  ||||   | |||| ||||||||| |
Db    323 TGGACGAGTGCTGAAGATAAGTGTATACGAAAACAATGTAAAACTCCTTCAGATCCTGAG 382

Qy    396 aatggcatggcacatgtgatcaaagacatccagttcggatcccaaattaaatattcttgt 455
          |||||| ||| ||||||   || || ||| ||||| |||||||  ||||| ||| |||||
Db    383 AATGGCTTGGTACATGTACACACAGGCATTCAGTTTGGATCCCGTATTAATTATACTTGT 442

Qy    456 cctaaaggataccgactcattggttcctcgtctgccacatgcatcatctcaggcaacact 515
            | |||||||||| |||||||||||||| |||||   |||  ||||| | |   | | |
Db    443 AATCAAGGATACCGCCTCATTGGTTCCTCCTCTGCTGTATGTGTCATCACTGATCAAAGT 502

Qy    516 gtcatttgggataataaaacacctgtttgtga 547
          ||   |||||||| | |  ||||| |||||||
Db    503 GTTGATTGGGATACTGAGGCACCTATTTGTGA 534


RESULT  13
AA107525
LOCUS       AA107525                614 bp    mRNA    linear   EST 04-NOV-1996
DEFINITION  mp08b07.r1 Life Tech mouse embryo 8 5dpc 10664019 Mus musculus cDNA
            clone IMAGE:568597 5' similar to gb:M23529 Mus musculus complement
            receptor (MOUSE);, mRNA sequence.
ACCESSION   AA107525
VERSION     AA107525.1  GI:1659424
```

```
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 614)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:343245
            Seq primer: -28M13 rev1 from Amersham
            High quality sequence stop: 431.
FEATURES             Location/Qualifiers
     source          1. .614
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:568597"
                     /clone_lib="Life Tech mouse embryo 8 5dpc 10664019"
                     /tissue_type="embryo"
                     /dev_stage="8.5dpc embryos"
                     /lab_host="DH10B"
                     /note="Organ: whole embryo; Vector: pCMV-SPORT2; Site_1:
                     SalI; Site_2: NotI; Cloned unidirectionally.  Primer:
                     Oligo dT.  8.5dpc embryos. pCMV-SPORT2 vector."
BASE COUNT      152 a    141 c    141 g    180 t
ORIGIN
```

```
  Query Match             33.9%;  Score 212.8;  DB 9;  Length 614;
  Best Local Similarity   71.4%;  Pred. No. 2.9e-48;
  Matches  280;  Conservative    0;  Mismatches  112;  Indels    0;  Gaps    0;

Qy    156 ctggtgttgctgctgtcctccttctccgatcaatgcaatgtcccggaatggcttccattt 215
          ||| ||||  ||||| | |      |  | ||| |||   | |||   |  |||||| | |
Db     96 CTGCTGTTCTTGCTGCCATTTACTTTGGGTCACTGCCCAGCCCCATCACAGCTTCCTTCT 155

Qy    216 gccaggcctaccaacctaactgatgactttgagtttcccattgggacatatctgaactat 275
          ||||  ||||  || ||||||||||| |    |||||||||||| |||||| ||   |||
Db    156 GCCAAACCTATAAATCTAACTGATGAATCCATGTTTCCCATTGGAACATATTTGTTGTAT 215

Qy    276 gaatgccgccctggttattccggaagaccgttttctatcatctgcctaaaaaactcagtc 335
          ||||| | ||| || |||   |   || | ||| ||||||| ||||  | || |||||  |
Db    216 GAATGTCTCCCAGGATATATCAAGAGGCAGTTCTCTATCACCTGCAAACAAGACTCAACC 275

Qy    336 tggacaagtgctaaggacaagtgcaaacgtaaatcatgtcgtaatcctccagatcctgtg 395
          ||||| |||||| | || ||||| | ||| |||  ||||   | |||| |||||||| |
Db    276 TGGACGAGTGCTGAAGATAAGTGTATACGAAAACAATGTAAAACTCCTTCAGATCCTGAG 335

Qy    396 aatggcatggcacatgtgatcaaagacatccagttcggatcccaaattaaatattcttgt 455
          |||||| ||| |||||||   || || ||| ||||| |||||||  ||||| ||| |||||
Db    336 AATGGCTTGGTACATGTACACACAGGCATTCAGTTTGGATCCCGTATTAATTATACTTGT 395

Qy    456 cctaaaggataccgactcattggttcctcgtctgccacatgcatcatctcaggcaacact 515
            | ||||||||||  |||||||||||||||  |||||   |||  ||||| | |   | | |
Db    396 AATCAAGGATACCGCCTCATTGGTTCCTCCTCTGCTGTATGTGTCATCACTGATCAAAGT 455

Qy    516 gtcatttgggataataaaacacctgtttgtga 547
          ||   |||||||| | |  ||||| |||||||
Db    456 GTTGATTGGGATACTGAGGCACCTATTTGTGA 487
```

```
RESULT  14
BI697900
LOCUS       BI697900                658 bp    mRNA    linear   EST 18-SEP-2001
DEFINITION  603344033F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5371665 5',
            mRNA sequence.
ACCESSION   BI697900
VERSION     BI697900.1  GI:15660529
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 658)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11946  row: a  column: 10
            High quality sequence stop: 658.
FEATURES             Location/Qualifiers
     source          1. .658
                     /organism="Mus musculus"
                     /strain="FVB/N-3"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5371665"
                     /clone_lib="NCI_CGAP_Mam2"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT      163 a    146 c    158 g    191 t
ORIGIN
```

```
  Query Match             33.9%;  Score 212.8;  DB 10;  Length 658;
  Best Local Similarity   71.4%;  Pred. No. 3e-48;
  Matches  280;  Conservative    0;  Mismatches  112;  Indels    0;  Gaps    0;

Qy    156 ctggtgttgctgctgtcctccttctccgatcaatgcaatgtcccggaatggcttccattt 215
          ||| ||||  ||||| | |      |  | ||| |||   | |||   |  |||||| | |
Db    123 CTGCTGTTCTTGCTGCCATTTACTTTGGGTCACTGCCCAGCCCCATCACAGCTTCCTTCT 182

Qy    216 gccaggcctaccaacctaactgatgactttgagtttcccattgggacatatctgaactat 275
          ||||  ||||  || ||||||||||| |    |||||||||||| |||||| ||   |||
Db    183 GCCAAACCTATAAATCTAACTGATGAATCCATGTTTCCCATTGGAACATATTTGTTGTAT 242

Qy    276 gaatgccgccctggttattccggaagaccgttttctatcatctgcctaaaaaactcagtc 335
          ||||| | ||| || |||   |   || | ||| ||||||| ||||  | || |||||  |
Db    243 GAATGTCTCCCAGGATATATCAAGAGGCAGTTCTCTATCACCTGCAAACAAGACTCAACC 302

Qy    336 tggacaagtgctaaggacaagtgcaaacgtaaatcatgtcgtaatcctccagatcctgtg 395
          ||||| |||||| | || ||||| | ||| |||  ||||   | |||| |||||||| |
Db    303 TGGACGAGTGCTGAAGATAAGTGTATACGAAAACAATGTAAAACTCCTTCAGATCCTGAG 362

Qy    396 aatggcatggcacatgtgatcaaagacatccagttcggatcccaaattaaatattcttgt 455
          |||||| ||| |||||||   || || ||| ||||| |||||||  ||||| ||| |||||
Db    363 AATGGCTTGGTACATGTACACACAGGCATTCAGTTTGGATCCCGTATTAATTATACTTGT 422

Qy    456 cctaaaggataccgactcattggttcctcgtctgccacatgcatcatctcaggcaacact 515
            | ||||||||||  |||||||||||||||  |||||   |||  ||||| | |   | | |
```

```
Db    423 AATCAAGGATACCGCCTCATTGGTTCCTCCTCTGCTGTATGTGTCATCACTGATCAAAGT 482

Qy    516 gtcatttgggataataaaacacctgtttgtga 547
          ||   |||||||| | |  ||||| |||||||
Db    483 GTTGATTGGGATACTGAGGCACCTATTTGTGA 514


RESULT  15
BI455761
LOCUS       BI455761                795 bp    mRNA    linear   EST 21-AUG-2001
DEFINITION  603171769F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5251289 5',
            mRNA sequence.
ACCESSION   BI455761
VERSION     BI455761.1  GI:15246417
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 795)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11634  row: i  column: 18
            High quality sequence stop: 795.
FEATURES             Location/Qualifiers
     source          1. .795
                     /organism="Mus musculus"
                     /strain="C57/B6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5251289"
                     /clone_lib="NCI_CGAP_Mam5"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="7 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Robin Humphreys,
                     NIH"
BASE COUNT      191 a    192 c    186 g    226 t
ORIGIN


  Query Match             33.9%;  Score 212.8;  DB 10;  Length 795;
  Best Local Similarity   71.4%;  Pred. No. 3.2e-48;
  Matches  280;  Conservative    0;  Mismatches  112;  Indels    0;  Gaps    0;

Qy    156 ctggtgttgctgctgtcctccttctccgatcaatgcaatgtcccggaatggcttccattt 215
          ||| ||||  ||||| | |     |  | ||| |||   | |||   |  |||||| | |
Db    161 CTGCTGTTCTTGCTGCCATTTACTTTGGGTCACTGCCCAGCCCCATCACAGCTTCCTTCT 220

Qy    216 gccaggcctaccaacctaactgatgactttgagtttcccattgggacatatctgaactat 275
          ||||  ||||  || ||||||||||| |    ||||||||||||| |||||| ||   |||
Db    221 GCCAAACCTATAAATCTAACTGATGAATCCATGTTTCCCATTGGAACATATTTGTTGTAT 280

Qy    276 gaatgccgccctggttattccggaagaccgttttctatcatctgcctaaaaaactcagtc 335
          ||||| | ||| || |||  |   || | ||| ||||||| ||||  | || |||||  |
Db    281 GAATGTCTCCCAGGATATATCAAGAGGCAGTTCTCTATCACCTGCAAACAAGACTCAACC 340

Qy    336 tggacaagtgctaaggacaagtgcaaacgtaaatcatgtcgtaatcctccagatcctgtg 395
          ||||| |||||| | || ||||| | ||| |||  ||||   | |||| |||||||||| |
Db    341 TGGACGAGTGCTGAAGATAAGTGTATACGAAAACAATGTAAAACTCCTTCAGATCCTGAG 400

Qy    396 aatggcatggcacatgtgatcaaagacatccagttcggatcccaaattaaatattcttgt 455
          |||||| ||| ||||||   || || ||| ||||| |||||||  ||||| ||| |||||
Db    401 AATGGCTTGGTACATGTACACACAGGCATTCAGTTTGGATCCCGTATTAATTATACTTGT 460

Qy    456 cctaaaggataccgactcattggttcctcgtctgccacatgcatcatctcaggcaacact 515
            | |||||||||| |||||||||||||| |||||   |||  ||||| | |   | | |
Db    461 AATCAAGGATACCGCCTCATTGGTTCCTCCTCTGCTGTATGTGTCATCACTGATCAAAGT 520

Qy    516 gtcatttgggataataaaacacctgtttgtga 547
          ||   |||||||| | |  ||||| |||||||
Db    521 GTTGATTGGGATACTGAGGCACCTATTTGTGA 552


Search completed: October  9, 2002, 17:37:05
Job time: 5552 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 9, 2002, 15:16:43 ; Search time 32.99 Seconds
(without alignments)
4668.451 Million cell updates/sec

Title: US-10-031-904-30
Perfect score: 627
Sequence: 1 cggactcagaagggacttcc.....ataataaaaatcttaaccga 627

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	352	56.1	6951	6	5256642-1	Patent No. 5256642
2	352	56.1	6951	6	5472939-1	Patent No. 5472939
3	267.4	42.6	605	2	US-08-769-967A-33	Sequence 33, Appl
4	218.6	34.9	5420	6	5256642-3	Patent No. 5256642
5	218.6	34.9	5420	6	5472939-3	Patent No. 5472939
6	156.4	24.9	1304	2	US-08-528-057-45	Sequence 45, Appl
7	148	23.6	1247	2	US-08-528-057-41	Sequence 41, Appl
8	118.4	18.9	1545	4	US-08-139-195-1	Sequence 1, Appli
9	117.6	18.8	1530	1	US-08-458-084-3	Sequence 3, Appli
10	117.6	18.8	1530	1	US-08-205-508-3	Sequence 3, Appli
11	117.6	18.8	1530	2	US-08-482-148-8	Sequence 8, Appli
12	117.6	18.8	1530	5	PCT-US95-02944-8	Sequence 8, Appli
13	117.6	18.8	1530	5	PCT-US95-02945-3	Sequence 3, Appli
14	117.6	18.8	1546	6	5514787-1	Patent No. 5514787
15	116.6	18.6	1878	2	US-08-435-149-17	Sequence 17, Appl
16	104.2	16.6	1986	2	US-08-528-057-43	Sequence 43, Appl
17	103.4	16.5	1659	2	US-08-528-057-1	Sequence 1, Appli
18	103.4	16.5	2124	1	US-08-087-007-1	Sequence 1, Appli
19	103.4	16.5	2124	3	US-08-483-433-1	Sequence 1, Appli
20	103.4	16.5	2124	5	PCT-US92-05920-1	Sequence 1, Appli
21	88.8	14.2	903	1	US-08-310-416A-11	Sequence 11, Appl
22	88.8	14.2	903	2	US-08-888-171-11	Sequence 11, Appl
23	88.8	14.2	1848	4	US-09-475-460A-31	Sequence 31, Appl
24	87.4	13.9	945	3	US-08-793-418-5	Sequence 5, Appli
25	87.4	13.9	1134	3	US-08-793-418-1	Sequence 1, Appli
26	87.4	13.9	1134	3	US-08-793-418-2	Sequence 2, Appli
27	87.4	13.9	1134	3	US-08-793-418-3	Sequence 3, Appli
28	87.4	13.9	1134	3	US-08-793-418-4	Sequence 4, Appli
29	82.4	13.1	101	2	US-08-356-361-5	Sequence 5, Appli
30	82.4	13.1	101	2	US-08-769-967A-5	Sequence 5, Appli
c 31	81.8	13.0	101	2	US-08-356-361-6	Sequence 6, Appli
c 32	81.8	13.0	101	2	US-08-769-967A-6	Sequence 6, Appli
33	75.4	12.0	101	2	US-08-356-361-3	Sequence 3, Appli
c 34	75.4	12.0	101	2	US-08-356-361-4	Sequence 4, Appli
35	75.4	12.0	101	2	US-08-769-967A-3	Sequence 3, Appli
c 36	75.4	12.0	101	2	US-08-769-967A-4	Sequence 4, Appli
37	60.8	9.7	1365	4	US-08-981-234B-1	Sequence 1, Appli
c 38	59	9.4	93	2	US-08-356-361-2	Sequence 2, Appli
c 39	59	9.4	93	2	US-08-769-967A-2	Sequence 2, Appli
40	58.4	9.3	94	2	US-08-356-361-7	Sequence 7, Appli
41	58.4	9.3	94	2	US-08-769-967A-7	Sequence 7, Appli
42	55.6	8.9	7218	1	US-08-232-463-14	Sequence 14, Appl
43	54.8	8.7	860	1	US-08-310-416A-19	Sequence 19, Appl
44	54.8	8.7	860	2	US-08-888-171-19	Sequence 19, Appl
45	54.8	8.7	969	1	US-08-310-416A-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
5256642-1
;Patent No. 5256642
; APPLICANT: FEARON, DOUGLAS T.;KLICKSTEIN, LLOYD B.;WONG,
;WINNIE W.;CARSON, GERALD R.;CONCINO, MICHAEL F.;IP, STEPHEN
;H.;MAKRIDES, SAVVAS;MARSH, HENRY C. JR.
; TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
;RECEPTOR 1 (CR1) AND A THROMBOLTIC AGENT, AND THE METHODS OF
;USE THEREOF
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/588,128
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 412,745
; FILING DATE: 26-SEP-1989
; APPLICATION NUMBER: 332,865
; FILING DATE: 03-APR-1989
; APPLICATION NUMBER: 176,532
; FILING DATE: 01-APR-1988
;SEQ ID NO:1:
; LENGTH: 6951
5256642-1

Query Match 56.1%; Score 352; DB 6; Length 6951;
Best Local Similarity 84.0%; Pred. No. 3e-103;
Matches 410; Conservative 0; Mismatches 75; Indels 3; Gaps 1;

```
Qy      66 acggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtccctttccttcc 125
           | |||| |  |  | ||    |   ||   || ||| | |    |   |||  ||   ||
Db      28 atgggggcctcttctccaagaagcccggagcctgtcgggccgccggcgcccggtctcccc 87

Qy     126 cggcgctttcctgggttgcttctggcggccctggtgttgct---gctgtcctccttctcc 182
               |||     || |  || |||||||   ||||| ||||   |||| |      ||
Db      88 ttctgctgcggaggatccctgctggcggttgtggtgctgcttgcgctgccggtggcctgg 147

Qy     183 gatcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgac 242
           | |||||||||| ||| ||||||||||||||||||||||||||||||||||||||||||
Db     148 ggtcaatgcaatgccccagaatggcttccatttgccaggcctaccaacctaactgatgag 207

Qy     243 tttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaaga 302
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208 tttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaaga 267

Qy     303 ccgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaa 362
           |||||||||||||||||||||||||||||||||||  |||||||||||| ||||| |
Db     268 ccgttttctatcatctgcctaaaaaactcagtctggactggtgctaaggacaggtgcaga 327

Qy     363 cgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagac 422
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||| |
Db    328 cgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggtgcatgtgatcaaaggc 387

Qy    423 atccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcc 482
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    388 atccagttcggatcccaaattaaatattcttgtactaaaggataccgactcattggttcc 447

Qy    483 tcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtt 542
          |||||||||||||||||||||||||  | |||||||||||||||||| |||||||| ||
Db    448 tcgtctgccacatgcatcatctcaggtgatactgtcatttgggataatgaaacacctatt 507

Qy    543 tgtgacag 550
          ||||||||
Db    508 tgtgacag 515


RESULT   2
5472939-1
;Patent No. 5472939
;    APPLICANT: FEARON, DOUGLAS T.;KLICKSTEIN, LLOYD B.;WONG,
;    JINNIE W.;CARSON, GERALD R.;CONCINO, MICHAEL F.;IP, STEPHEN
;    Y.;MAKRIDES, SAVVAS;MARSH, HENRY C. JR.
;    TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
;MEDIATED DISORDERS
;    NUMBER OF SEQUENCES: 30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/138,825
;      FILING DATE: 19-OCT-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 588,128
;      FILING DATE: 24-SEP-1990
;      APPLICATION NUMBER: 412,745
;      FILING DATE: 26-SEP-1989
;      APPLICATION NUMBER: 332,865
;      FILING DATE: 03-APR-1989
;      APPLICATION NUMBER: 176,532
;      FILING DATE: 01-APR-1988
;SEQ ID NO:1:
;      LENGTH: 6951
5472939-1


  Query Match             56.1%;  Score 352;  DB 6;  Length 6951;
  Best Local Similarity   84.0%;  Pred. No. 3e-103;
  Matches  410;  Conservative    0;  Mismatches   75;  Indels    3;  Gaps    1;

Qy     66 acggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtccctttccttcc 125
          | |||| |  |  | ||    |   ||  || ||| | |    |   |||  ||   ||
       28 atgggggcctcttctccaagaagcccggagcctgtcgggccgccggcgcccggtctcccc 87

Qy    126 cggcgctttcctgggttgcttctggcggccctggtgttgct---gctgtcctccttctcc 182
            |||     || |  || |||||||   ||||| ||||   |||| |      ||
Db     88 ttctgctgcggaggatccctgctggcggttgtggtgctgcttgcgctgccggtggcctgg 147

Qy    183 gatcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgac 242
          | |||||||||| ||| ||||||||||||||||||||||||||||||||||||||||||
Db    148 ggtcaatgcaatgccccagaatggcttccatttgccaggcctaccaacctaactgatgag 207

Qy    243 tttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaaga 302
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    208 tttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaaga 267

Qy    303 ccgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaa 362
          |||||||||||||||||||||||||||||||||||  |||||||||||| ||||| |
Db    268 ccgttttctatcatctgcctaaaaaactcagtctggactggtgctaaggacaggtgcaga 327

Qy    363 cgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagac 422
          |||||||||||||||||||||||||||||||||||||||||||  |||||||||||| |
Db    328 cgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggtgcatgtgatcaaaggc 387

Qy    423 atccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcc 482
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    388 atccagttcggatcccaaattaaatattcttgtactaaaggataccgactcattggttcc 447

Qy    483 tcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtt 542
          |||||||||||||||||||||||||  | |||||||||||||||||| |||||||| ||
Db    448 tcgtctgccacatgcatcatctcaggtgatactgtcatttgggataatgaaacacctatt 507

Qy    543 tgtgacag 550
          ||||||||
Db    508 tgtgacag 515


RESULT   3
US-08-769-967A-33
; Sequence 33, Application US/08769967A
; Patent No. 5859223
;  GENERAL INFORMATION:
;    APPLICANT:  Mossakowska, Danuta E.I.
;    APPLICANT:  Smith, Richard A.G.
;    APPLICANT:  Dodd, Ian
;    APPLICANT:  Freeman, Anne Mary
;    TITLE OF INVENTION:  Soluble CR1 Derivatives
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporate Intellectual Property
;      STREET:  P.O. Box 1539
;      CITY:  King of Prussia
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19406
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/769,967A
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/440,569
;      FILING DATE:  15-May-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  King, William T.
;      REGISTRATION NUMBER:  30,954
;      REFERENCE/DOCKET NUMBER:  P30423C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610) 270-5364
;      TELEFAX:  (610) 270-5090
;  INFORMATION FOR SEQ ID NO:  33:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  605 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-769-967A-33


  Query Match             42.6%;  Score 267.4;  DB 2;  Length 605;
  Best Local Similarity   83.3%;  Pred. No. 1.4e-76;
  Matches  304;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy    186 caatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatgacttt 245
          || ||||| |  |||||||||||| || || ||  | || |||||||| |||||||| |||
Db      7 CAGTGCAACGCTCCGGAATGGCTGCCGTTCGCGCGCCCGACCAACCTGACTGATGAATTT 66

Qy    246 gagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaagaccg 305
          ||||| || || || || || |||||||| |||||||||| |||||  |||  | |||
Db     67 GAGTTCCCGATCGGTACCTACCTGAACTACGAATGCCGCCCGGGTTATAGCGGCCGCCCG 126

Qy    306 ttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgcaaacgt 365
          ||||||||||||||| |||||||| ||||||||  ||||||||||   |||  ||||
Db    127 TTTTCTATCATCTGCCTGAAAAACTCTGTCTGGACTGGTGCTAAGGACCGTTGCCGACGT 186
```

```
Qy    366 aaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaagacatc 425
          ||||| |||||||||||| |||||||| || || |||||||  |||||||||||| ||||
Db    187 AAATCTTGTCGTAATCCGCCAGATCCGGTTAACGGCATGGTGCATGTGATCAAAGGCATC 246

Qy    426 cagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggttcctcg 485
          |||||||| ||||||||||||||||||||| ||||||| ||||| || ||||||||||
Db    247 CAGTTCGGTTCCCAAATTAAATATTCTTGTACTAAAGGTTACCGTCTGATTGGTTCCTCC 306

Qy    486 tctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctgtttgt 545
            || |||||||||||||| ||  | |||||||||||||||||| |||||||  |||||
Db    307 AGCGCTACATGCATCATCTCTGGTGATACTGTCATTTGGGATAATGAAACACCGATTTGT 366

Qy    546 gacag 550
          ||| |
Db    367 GACCG 371


RESULT   4
5256642-3
;Patent No. 5256642
;    APPLICANT: FEARON, DOUGLAS T.;KLICKSTEIN, LLOYD B.;WONG,
;WINNIE W.;CARSON, GERALD R.;CONCINO, MICHAEL F.;IP, STEPHEN
;H.;MAKRIDES, SAVVAS;MARSH, HENRY C. JR.
;    TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
;RECEPTOR 1 (CR1) AND A THROMBOLTIC AGENT, AND THE METHODS OF
;USE THEREOF
;    NUMBER OF SEQUENCES: 30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/588,128
;      FILING DATE: 24-SEP-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 412,745
;      FILING DATE: 26-SEP-1989
;      APPLICATION NUMBER: 332,865
;      FILING DATE: 03-APR-1989
;      APPLICATION NUMBER: 176,532
;      FILING DATE: 01-APR-1988
;SEQ ID NO:3:
;      LENGTH: 5420
5256642-3


  Query Match             34.9%;  Score 218.6;  DB 6;  Length 5420;
  Best Local Similarity   74.5%;  Pred. No. 2.6e-60;
  Matches  275;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy     181 ccgatcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatg 240
           | | ||| || ||   ||| ||   | ||||||||||||| ||||| | || || | |||
Db    2674 ctggtcactgtaaaaccccagagcagtttccatttgccagtcctacgatcccaattaatg 2733

Qy     241 actttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaa 300
           |||||||||||||  | ||||||| | |||| |||||||||||| ||||| ||||  || |
Db    2734 actttgagtttccagtcgggacatctttgaattatgaatgccgtcctgggtattttggga 2793

Qy     301 gaccgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgca 360
           |  ||| ||||||  ||||||| ||||||  |||||| |||||| | | ||||| || |
Db    2794 aaatgttctctatctcctgcctagaaaacttggtctggtcaagtgttgaagacaactgta 2853

Qy     361 aacgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaag 420
           ||| |||||||||  |   ||| ||||| ||  | ||||| |||  ||| | | || ||
Db    2854 gacgaaaatcatgtggacctccaccagaacccttcaatggaatggtgcatataaacacag 2913

Qy     421 acatccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggtt 480
           | |  ||||| |||||   | |||| |||||||||  | |||| |  ||||||||||||
Db    2914 atacacagtttggatcaacagttaattattcttgtaatgaagggtttcgactcattggtt 2973

Qy     481 cctcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctg 540
           || | ||| | || ||  || |||||||||| | |||||  |||||||| ||  |||||
Db    2974 ccccatctactacttgtctcgtctcaggcaataatgtcacatgggataagaaggcaccta 3033

Qy     541 tttgtgaca 549
           |||||||
Db    3034 tttgtgaga 3042


RESULT   5
5472939-3
;Patent No. 5472939
;    APPLICANT: FEARON, DOUGLAS T.;KLICKSTEIN, LLOYD B.;WONG,
;WINNIE W.;CARSON, GERALD R.;CONCINO, MICHAEL F.;IP, STEPHEN
;H.;MAKRIDES, SAVVAS;MARSH, HENRY C. JR.
;    TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
;MEDIATED DISORDERS
;    NUMBER OF SEQUENCES: 30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/138,825
;      FILING DATE: 19-OCT-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 588,128
;      FILING DATE: 24-SEP-1990
;      APPLICATION NUMBER: 412,745
;      FILING DATE: 26-SEP-1989
;      APPLICATION NUMBER: 332,865
;      FILING DATE: 03-APR-1989
;      APPLICATION NUMBER: 176,532
;      FILING DATE: 01-APR-1988
;SEQ ID NO:3:
;      LENGTH: 5420
5472939-3


  Query Match             34.9%;  Score 218.6;  DB 6;  Length 5420;
  Best Local Similarity   74.5%;  Pred. No. 2.6e-60;
  Matches  275;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy     181 ccgatcaatgcaatgtcccggaatggcttccatttgccaggcctaccaacctaactgatg 240
           | | ||| || ||   ||| ||   | ||||||||||||| ||||| | || || | |||
Db    2674 ctggtcactgtaaaaccccagagcagtttccatttgccagtcctacgatcccaattaatg 2733

Qy     241 actttgagtttcccattgggacatatctgaactatgaatgccgccctggttattccggaa 300
           |||||||||||||  | ||||||| | |||| |||||||||||| ||||| ||||  || |
Db    2734 actttgagtttccagtcgggacatctttgaattatgaatgccgtcctgggtattttggga 2793

Qy     301 gaccgttttctatcatctgcctaaaaaactcagtctggacaagtgctaaggacaagtgca 360
           |  ||| ||||||  ||||||| ||||||  |||||| |||||| | | ||||| || |
Db    2794 aaatgttctctatctcctgcctagaaaacttggtctggtcaagtgttgaagacaactgta 2853

Qy     361 aacgtaaatcatgtcgtaatcctccagatcctgtgaatggcatggcacatgtgatcaaag 420
           ||| |||||||||  |   ||| ||||| ||  | ||||| |||  ||| | | || ||
Db    2854 gacgaaaatcatgtggacctccaccagaacccttcaatggaatggtgcatataaacacag 2913

Qy     421 acatccagttcggatcccaaattaaatattcttgtcctaaaggataccgactcattggtt 480
           | |  ||||| |||||   | |||| |||||||||  | |||| |  ||||||||||||
Db    2914 atacacagtttggatcaacagttaattattcttgtaatgaagggtttcgactcattggtt 2973

Qy     481 cctcgtctgccacatgcatcatctcaggcaacactgtcatttgggataataaaacacctg 540
           || | ||| | || ||  || |||||||||| | |||||  |||||||| ||  |||||
Db    2974 ccccatctactacttgtctcgtctcaggcaataatgtcacatgggataagaaggcaccta 3033

Qy     541 tttgtgaca 549
           |||||||
Db    3034 tttgtgaga 3042


RESULT   6
US-08-528-057-45
; Sequence 45, Application US/08528057
; Patent No. 5846715
;  GENERAL INFORMATION:
;    APPLICANT:  PURCELL, Damian F. J.
;    APPLICANT:  RUSSELL, Sarah M.
;    APPLICANT:  McKENZIE, Ian F. C.
;    TITLE OF INVENTION:  CD46 VARIANTS
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
```

```
;       STREET:  3000 K Street, N.W., Suite 500
;       CITY:  Washington
;       STATE:  D.C.
;       COUNTRY:  USA
;       ZIP:  20007-5109
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/528,057
;       FILING DATE:  CONCURRENTLY HEREWITH
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/961,686
;       FILING DATE:  11-JAN-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/AU91/00199
;       FILING DATE:  10-MAY-1991
      PRIOR APPLICATION DATA:
        APPLICATION NUMBER:  AU PK0133/90
        FILING DATE:  11-MAY-1990
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  BENT, Stephen A.
;       REGISTRATION NUMBER:  29,768
;       REFERENCE/DOCKET NUMBER:  17227/112 DACO
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202)672-5300
;       TELEFAX:  (202)672-5399
;       TELEX:  904136
;   INFORMATION FOR SEQ ID NO:  45:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1304 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  94..1065
;     FEATURE:
;       NAME/KEY:  mat_peptide
;       LOCATION:  196..1065
US-08-528-057-45


  Query Match             24.9%;  Score 156.4;  DB 2;  Length 1304;
  Best Local Similarity   86.9%;  Pred. No. 1.3e-40;
  Matches  172;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       2 ggactcagaagggacttccctgctcggctggctttcggtttctctgctcacctccggata 61
           |||| |||||||||||||||||||||||||| ||||||||||||||  |||||||| |
Db       9 GGACCCAGAAGGGACTTCCCTGCTCGGCTGGCTCTCGGTTTCTCTGCTTTCCTCCGGAGA 68

Qy      62 aatcacggggtctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtccctttcc 121
           ||| || | |||| |||||||| ||||| |||||||| ||| | |||| ||||||||||
Db      69 AATAACAGCGTCTTCCGCGCCGCGCATGGAGCCTCCCGGCCGCCGCGAGTGTCCCTTTCC 128

Qy     122 ttcccggcgctttcctgggttgcttctggcggccctggtgttgctgctgtcctccttctc 181
           |||| |||||||||||||||||||||||||||| |||||||||||||| |||||||||
Db     129 TTCCTGGCGCTTTCCTGGGTTGCTTCTGGCGGCCATGGTGTTGCTGCTGTACTCCTTCTC 188

Qy     182 cgatcaatgcaatgtccc 199
           ||||   ||  | |  ||
Db     189 CGATGCCTGTGAGGAGCC 206


RESULT   7
US-08-528-057-41
; Sequence 41, Application US/08528057
; Patent No. 5846715
;  GENERAL INFORMATION:
;    APPLICANT:  PURCELL, Damian F. J.
;    APPLICANT:  RUSSELL, Sarah M.
;    APPLICANT:  McKENZIE, Ian F. C.
;    TITLE OF INVENTION:  CD46 VARIANTS
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/528,057
;      FILING DATE:  CONCURRENTLY HEREWITH
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/961,686
;      FILING DATE:  11-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/AU91/00199
;      FILING DATE:  10-MAY-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PK0133/90
;      FILING DATE:  11-MAY-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  17227/112 DACO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  41:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1247 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  83..1192
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  185..1192
US-08-528-057-41


  Query Match             23.6%;  Score 148;  DB 2;  Length 1247;
  Best Local Similarity   86.7%;  Pred. No. 6.4e-38;
  Matches  163;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy      12 gggacttccctgctcggctggctttcggtttctctgctcacctccggataaatcacgggg 71
           |||||||||||||||||||||||| ||||||||||||||  |||||||| |||| || | |
Db       8 GGGACTTCCCTGCTCGGCTGGCTCTCGGTTTCTCTGCTTTCCTCCGGAGAAATAACAGCG 67

Qy      72 tctcccgcgccgctcatggcgcctcccgtccgtctcgagcgtccctttccttcccggcgc 131
           ||| |||||||||| ||||| |||||||| ||| | |||| ||||||||||||||| |||||
Db      68 TCTTCCGCGCCGCGCATGGAGCCTCCCGGCCGCCGCGAGTGTCCCTTTCCTTCCTGGCGC 127

Qy     132 tttcctgggttgcttctggcggccctggtgttgctgctgtcctccttctccgatcaatgc 191
           ||||||||||||||||||||||||| |||||||||||||| ||||||||||||   ||
Db     128 TTTCCTGGGTTGCTTCTGGCGGCCATGGTGTTGCTGCTGTACTCCTTCTCCGATGCCTGT 187

Qy     192 aatgtccc 199
             | |  ||
Db     188 GAGGAGCC 195
```

```
RESULT   8
US-08-139-195-1
; Sequence 1, Application US/08139195
; Patent No. 6218520
;  GENERAL INFORMATION:
;    APPLICANT:  Atkinson, John P.
;    TITLE OF INVENTION:  RECOMBINANTLY PRODUCED HUMAN MEMBRANE
;    TITLE OF INVENTION:  COFACTOR PROTEIN (MCP)
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  1100 Peachtree Street, Suite 2800
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-4530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/139,195
;      FILING DATE:  20-OCT-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/948,350
;      FILING DATE:  21-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/384,210
;      FILING DATE:  21-JUL-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  WU100CON
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404)-815-6508
;      TELEFAX:  (404)-815-6555
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1545 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Human t-Cell Line HSB2
US-08-139-195-1


  Query Match            18.9%;  Score 118.4;  DB 4;  Length 1545;
  Best Local Similarity  83.8%;  Pred. No. 2.5e-28;
  Matches  134;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy      44 tctgctcacctccggataaatcacggggtctcccgcgccgctcatggcgcctcccgtccg 103
           ||||||  |||||||| |||| || | |||| ||||||||| ||||| ||||||||
Db       1 TCTGCTTTCCTCCGGAGAAATAACAGCGTCTTCCGCGCCGCGCATGGAGCCTCCCGGCCG 60

Qy     104 tctcgagcgtccctttccttcccggcgctttcctgggttgcttctggcggccctggtgtt 163
             | |||| |||||||||||||| ||||||||||||||||||||||||||||| |||||||
Db      61 CCGCGAGTGTCCCTTTCCTTCCTGGCGCTTTCCTGGGTTGCTTCTGGCGGCCATGGTGTT 120

Qy     164 gctgctgtcctccttctccgatcaatgcaatgtcccggaa 203
           ||||||||| ||||||||||||   ||  | |  ||  ||
Db     121 GCTGCTGTACTCCTTCTCCGATGCCTGTGAGGAGCCACAA 160


RESULT   9
US-08-458-084-3
; Sequence 3, Application US/08458084
; Patent No. 5624837
;  GENERAL INFORMATION:
;    APPLICANT:  Fodor, William L
;    APPLICANT:  Rollins, Scott
;    APPLICANT:  Squinto, Stephen P
;    TITLE OF INVENTION:  Chimeric Complement
;    TITLE OF INVENTION:  Inhibitor Proteins
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maurice M. Klee
;      STREET:  1951 Burr Street
;      CITY:  Fairfield
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06430
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 750 Kb storage
;      COMPUTER:  Dell 486/50
;      OPERATING SYSTEM:  DOS 6.2
;      SOFTWARE:  WordPerfect 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,084
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Klee, Maurice M.
;      REGISTRATION NUMBER:  30,399
;      REFERENCE/DOCKET NUMBER:  ALX-120
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203) 255-1400
;      TELEFAX:  (203) 254-1101
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1530 bases
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Double
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  cDNA to mRNA
;      DESCRIPTION:  MCP (CD46) full length cDNA
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    PUBLICATION INFORMATION:
;      AUTHORS:  Lublin, D.M.
;      AUTHORS:  Liszewski,M.K.
;      AUTHORS:  Post, T.W.
;      AUTHORS:  Arce, M.A.
;      AUTHORS:  LeBeau, M.M.
;      AUTHORS:  Rebentisch, M.B.
;      AUTHORS:  Lemons, R.S.
;      AUTHORS:  Seya, T.
;      AUTHORS:  Atkinson, J.P.
;      TITLE:  Molecular cloning and Chromosomal
;      TITLE:  Localization of Membrane Cofactor
;      TITLE:  Protein (MCP): Evidence for Inclusion
;      TITLE:  in the Multi-Gene Family of
;      TITLE:  Complement-Regulatory Proteins.
;      JOURNAL:  Journal of Experimental Medicine
;      VOLUME:  168
;      PAGES:  181-194
;      DATE:  1988
US-08-458-084-3


  Query Match            18.8%;  Score 117.6;  DB 1;  Length 1530;
  Best Local Similarity  84.6%;  Pred. No. 4.5e-28;
  Matches  132;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy      44 tctgctcacctccggataaatcacggggtctcccgcgccgctcatggcgcctcccgtccg 103
           ||||||  |||||||| |||| || | |||| ||||||||| ||||| ||||||||
Db       1 TCTGCTTTCCTCCGGAGAAATAACAGCGTCTTCCGCGCCGCGCATGGAGCCTCCCGGCCG 60

Qy     104 tctcgagcgtccctttccttcccggcgctttcctgggttgcttctggcggccctggtgtt 163
```

```
            | ||||  ||||||||||||||| |||||||||||||||||||||||||||| |||||||
Db      61 CCGCGAGTGTCCCTTTCCTTCCTGGCGCTTTCCTGGGTTGCTTCTGGCGGCCATGGTGTT 120

Qy     164 gctgctgtcctccttctccgatcaatgcaatgtccc 199
           |||||||| ||||||||||||   ||  | |  ||
Db     121 GCTGCTGTACTCCTTCTCCGATGCCTGTGAGGAGCC 156


RESULT  10
US-08-205-508-3
; Sequence 3, Application US/08205508
; Patent No. 5627264
;  GENERAL INFORMATION:
;    APPLICANT:  Fodor, William L
;    APPLICANT:  Rollins, Scott
;    APPLICANT:  Squinto, Stephen P
;    TITLE OF INVENTION:  Chimeric Complement
;    TITLE OF INVENTION:  Inhibitor Proteins
;    NUMBER OF SEQUENCES:  19
     CORRESPONDENCE ADDRESS:
       ADDRESSEE:  Maurice M. Klee
       STREET:  1951 Burr Street
;      CITY:  Fairfield
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06430
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 750 Kb storage
;      COMPUTER:  Dell 486/50
;      OPERATING SYSTEM:  DOS 6.2
;      SOFTWARE:  WordPerfect 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/205,508
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Klee, Maurice M.
;      REGISTRATION NUMBER:  30,399
;      REFERENCE/DOCKET NUMBER:  ALX-120
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203) 255-1400
;      TELEFAX:  (203) 254-1101
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1530 bases
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Double
       TOPOLOGY:  Linear
     MOLECULE TYPE:  cDNA to mRNA
       DESCRIPTION:  MCP (CD46) full length cDNA
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    PUBLICATION INFORMATION:
;      AUTHORS:  Lublin, D.M.
;      AUTHORS:  Liszewski,M.K.
;      AUTHORS:  Post, T.W.
;      AUTHORS:  Arce, M.A.
;      AUTHORS:  LeBeau, M.M.
;      AUTHORS:  Rebentisch, M.B.
;      AUTHORS:  Lemons, R.S.
;      AUTHORS:  Seya, T.
;      AUTHORS:  Atkinson, J.P.
;      TITLE:  Molecular cloning and Chromosomal
;      TITLE:  Localization of Membrane Cofactor
;      TITLE:  Protein (MCP): Evidence for Inclusion
;      TITLE:  in the Multi-Gene Family of
;      TITLE:  Complement-Regulatory Proteins.
;      JOURNAL:  Journal of Experimental Medicine
;      VOLUME:  168
;      PAGES:  181-194
;      DATE:  1988
```

```
US-08-205-508-3


  Query Match            18.8%;  Score 117.6;  DB 1;  Length 1530;
  Best Local Similarity  84.6%;  Pred. No. 4.5e-28;
  Matches  132;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy      44 tctgctcacctccggataaatcacggggtctcccgcgccgctcatggcgcctcccgtccg 103
           ||||||  |||||||| |||| || | |||| ||||||||| ||||| |||||||| |||
Db       1 TCTGCTTTCCTCCGGAGAAATAACAGCGTCTTCCGCGCCGCGCATGGAGCCTCCCGGCCG 60

Qy     104 tctcgagcgtccctttccttcccggcgctttcctgggttgcttctggcggccctggtgtt 163
           | |||| |||||||||||||| |||||||||||||||||||||||||||||| |||||||
Db      61 CCGCGAGTGTCCCTTTCCTTCCTGGCGCTTTCCTGGGTTGCTTCTGGCGGCCATGGTGTT 120

Qy     164 gctgctgtcctccttctccgatcaatgcaatgtccc 199
           |||||||| ||||||||||||   ||  | |  ||
Db     121 GCTGCTGTACTCCTTCTCCGATGCCTGTGAGGAGCC 156


RESULT  11
US-08-482-148-8
; Sequence 8, Application US/08482148
; Patent No. 5847082
;  GENERAL INFORMATION:
;    APPLICANT:  Rother, Russell
;    APPLICANT:  Rollins, Scott
;    APPLICANT:  Squinto, Stephen P
;    TITLE OF INVENTION:  Terminal Complement
;    TITLE OF INVENTION:  Inhibitor Fusion Genes and Proteins
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seth A. Fidel
;      STREET:  Alexion Pharmaceuticals, 25 Science Park
;      CITY:  New Haven
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06511
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.4 Mb storage
;      COMPUTER:  IBM compatible (Pentium)
;      OPERATING SYSTEM:  Windows 95 under MS DOS
;      SOFTWARE:  Microsoft Word for Windows
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,148
;      FILING DATE:  6/07/95
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fidel, Seth A.
;      REGISTRATION NUMBER:  38,449
;      REFERENCE/DOCKET NUMBER:  ALX-129.1 Div
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203) 776 1790
;      TELEFAX:  (203) 772 3655
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1530 base pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Double
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  cDNA to mRNA
;      DESCRIPTION:  MCP (CD46) full length cDNA
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    PUBLICATION INFORMATION:
;      AUTHORS:  Lublin, D.M.
;      AUTHORS:  Liszewski,M.K.
;      AUTHORS:  Post, T.W.
;      AUTHORS:  Arce, M.A.
;      AUTHORS:  LeBeau, M.M.
;      AUTHORS:  Rebentisch, M.B.
```

```
;       AUTHORS:  Lemons, R.S.
;       AUTHORS:  Seya, T.
;       AUTHORS:  Atkinson, J.P.
;       TITLE:  Molecular cloning and Chromosomal
;       TITLE:  Localization of Membrane Cofactor
;       TITLE:  Protein (MCP): Evidence for Inclusion
;       TITLE:  in the Multi-Gene Family of
;       TITLE:  Complement-Regulatory Proteins.
;       JOURNAL:  Journal of Experimental Medicine
;       VOLUME:  168
;       PAGES:  181-194
;       DATE:  1988
US-08-482-148-8


  Query Match             18.8%;  Score 117.6;  DB 2;  Length 1530;
  Best Local Similarity   84.6%;  Pred. No. 4.5e-28;
  Matches  132;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy      44 tctgctcacctccggataaatcacggggtctcccgcgccgctcatggcgcctcccgtccg 103
           ||||||  |||||||| |||| || | |||| |||||||||| ||||| |||||||| |||
Db       1 TCTGCTTTCCTCCGGAGAAATAACAGCGTCTTCCGCGCCGCGCATGGAGCCTCCCGGCCG 60

Qy     104 tctcgagcgtccctttccttcccggcgctttcctgggttgcttctggcggccctggtgtt 163
           | |||| |||||||||||||| ||||||||||||||||||||||||||||| |||||||
Db      61 CCGCGAGTGTCCCTTTCCTTCCTGGCGCTTTCCTGGGTTGCTTCTGGCGGCCATGGTGTT 120

Qy     164 gctgctgtcctccttctccgatcaatgcaatgtccc 199
           |||||||| ||||||||||||   ||  | |  ||
Db     121 GCTGCTGTACTCCTTCTCCGATGCCTGTGAGGAGCC 156


RESULT  12
PCT-US95-02944-8
; Sequence 8, Application PC/TUS9502944
;  GENERAL INFORMATION:
;    APPLICANT:  Rother, Russell
;    APPLICANT:  Rollins, Scott
;    APPLICANT:  Squinto, Stephen P
;    TITLE OF INVENTION:  Terminal Complement
;    TITLE OF INVENTION:  Inhibitor Fusion Genes and Proteins
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maurice M. Klee
;      STREET:  1951 Burr Street
;      CITY:  Fairfield
;      STATE:  Connecticut
       COUNTRY:  USA
       ZIP:  06430
     COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 720 Kb storage
;      COMPUTER:  Dell 486/50
;      OPERATING SYSTEM:  DOS 6.2
;      SOFTWARE:  Word Perfect 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/02944
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/205,720
;      FILING DATE:  3-MAR-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Klee, Maurice M.
;      REGISTRATION NUMBER:  30,399
;      REFERENCE/DOCKET NUMBER:  ALX-129PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203) 255 1400
;      TELEFAX:  (203) 254 1101
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1530 base pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Double
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  cDNA to mRNA
;      DESCRIPTION:  MCP (CD46) full length cDNA
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    PUBLICATION INFORMATION:
;      AUTHORS:  Lublin, D.M.
;      AUTHORS:  Liszewski,M.K.
;      AUTHORS:  Post, T.W.
;      AUTHORS:  Arce, M.A.
;      AUTHORS:  LeBeau, M.M.
;      AUTHORS:  Rebentisch, M.B.
;      AUTHORS:  Lemons, R.S.
;      AUTHORS:  Seya, T.
;      AUTHORS:  Atkinson, J.P.
;      TITLE:  Molecular cloning and Chromosomal
;      TITLE:  Localization of Membrane Cofactor
;      TITLE:  Protein (MCP): Evidence for
;      TITLE:  Inclusion in the Multi-Gene Family
;      TITLE:  of Complement-Regulatory Proteins.
;      JOURNAL:  Journal of Experimental Medicine
;      VOLUME:  168
;      PAGES:  181-194
;      DATE:  1988
PCT-US95-02944-8


  Query Match             18.8%;  Score 117.6;  DB 5;  Length 1530;
  Best Local Similarity   84.6%;  Pred. No. 4.5e-28;
  Matches  132;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy      44 tctgctcacctccggataaatcacggggtctcccgcgccgctcatggcgcctcccgtccg 103
           ||||||  |||||||| |||| || | |||| |||||||||| ||||| |||||||| |||
Db       1 TCTGCTTTCCTCCGGAGAAATAACAGCGTCTTCCGCGCCGCGCATGGAGCCTCCCGGCCG 60

Qy     104 tctcgagcgtccctttccttcccggcgctttcctgggttgcttctggcggccctggtgtt 163
           | |||| |||||||||||||| ||||||||||||||||||||||||||||| |||||||
Db      61 CCGCGAGTGTCCCTTTCCTTCCTGGCGCTTTCCTGGGTTGCTTCTGGCGGCCATGGTGTT 120

Qy     164 gctgctgtcctccttctccgatcaatgcaatgtccc 199
           |||||||| ||||||||||||   ||  | |  ||
Db     121 GCTGCTGTACTCCTTCTCCGATGCCTGTGAGGAGCC 156


RESULT  13
PCT-US95-02945-3
; Sequence 3, Application PC/TUS9502945
;  GENERAL INFORMATION:
;    APPLICANT:  Fodor, William L
;    APPLICANT:  Rollins, Scott
;    APPLICANT:  Squinto, Stephen P
;    TITLE OF INVENTION:  Chimeric Complement
;    TITLE OF INVENTION:  Inhibitor Proteins
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Maurice M. Klee
;      STREET:  1951 Burr Street
;      CITY:  Fairfield
;      STATE:  Connecticut
;      COUNTRY:  USA
;      ZIP:  06430
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 750 Kb storage
;      COMPUTER:  Dell 486/50
;      OPERATING SYSTEM:  DOS 6.2
;      SOFTWARE:  WordPerfect 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/02945
;      FILING DATE:
;      CLASSIFICATION:
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/205,508
;      FILING DATE:  3-MAR-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Klee, Maurice M.
;      REGISTRATION NUMBER:  30,399
;      REFERENCE/DOCKET NUMBER:  ALX-120PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203) 255-1400
;      TELEFAX:  (203) 254-1101
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1530 bases
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Double
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  cDNA to mRNA
;      DESCRIPTION:  MCP (CD46) full length cDNA
     HYPOTHETICAL:  No
     ANTI-SENSE:  No
     ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    PUBLICATION INFORMATION:
;      AUTHORS:  Lublin, D.M.
;      AUTHORS:  Liszewski,M.K.
;      AUTHORS:  Post, T.W.
;      AUTHORS:  Arce, M.A.
;      AUTHORS:  LeBeau, M.M.
;      AUTHORS:  Rebentisch, M.B.
;      AUTHORS:  Lemons, R.S.
;      AUTHORS:  Seya, T.
;      AUTHORS:  Atkinson, J.P.
;      TITLE:  Molecular cloning and Chromosomal
;      TITLE:  Localization of Membrane Cofactor
;      TITLE:  Protein (MCP): Evidence for
;      TITLE:  Inclusion in the Multi-Gene Family
;      TITLE:  of Complement-Regulatory Proteins.
;      JOURNAL:  Journal of Experimental Medicine
;      VOLUME:  168
;      PAGES:  181-194
;      DATE:  1988
PCT-US95-02945-3


  Query Match            18.8%;  Score 117.6;  DB 5;  Length 1530;
  Best Local Similarity   84.6%;  Pred. No. 4.5e-28;
  Matches  132;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy      44 tctgctcacctccggataaatcacggggtctcccgcgccgctcatggcgcctcccgtccg 103
           ||||||  |||||||| |||| || | |||| ||||||||| ||||| |||||||| |||
Db       1 TCTGCTTTCCTCCGGAGAAATAACAGCGTCTTCCGCGCCGCGCATGGAGCCTCCCGGCCG 60

Qy     104 tctcgagcgtccctttccttcccggcgctttcctgggttgcttctggcggccctggtgtt 163
             | |||| |||||||||||||| ||||||||||||||||||||||||||||| |||||||
Db      61 CCGCGAGTGTCCCTTTCCTTCCTGGCGCTTTCCTGGGTTGCTTCTGGCGGCCATGGTGTT 120

Qy     164 gctgctgtcctccttctccgatcaatgcaatgtccc 199
           ||||||||| ||||||||||||||   ||  | |  ||
Db     121 GCTGCTGTACTCCTTCTCCGATGCCTGTGAGGAGCC 156


RESULT  14
5514787-1
;Patent No. 5514787
;    APPLICANT: ATKINSON, JOHN P.
;    TITLE OF INVENTION: DNA SEQUENCES ENCODING HUMAN MEMBRANE
;COFACTOR PROTEIN (MCP)
;    NUMBER OF SEQUENCES: 2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/948,350
;      FILING DATE: 21-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 384,210
;      FILING DATE: 21-JUL-1989
;SEQ ID NO:1:
;      LENGTH: 1546
5514787-1


  Query Match            18.8%;  Score 117.6;  DB 6;  Length 1546;
  Best Local Similarity   84.6%;  Pred. No. 4.5e-28;
  Matches  132;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy      44 tctgctcacctccggataaatcacggggtctcccgcgccgctcatggcgcctcccgtccg 103
           ||||||  |||||||| |||| || | |||| ||||||||| ||||| |||||||| |||
Db       1 tctgctttcctccggagaaataacagcgtcttccgcgccgcgcatggagcctcccggccg 60

Qy     104 tctcgagcgtccctttccttcccggcgctttcctgggttgcttctggcggccctggtgtt 163
             | |||| |||||||||||||| ||||||||||||||||||||||||||||| |||||||
Db      61 ccgcgagtgtccctttccttcctggcgctttcctgggttgcttctggcggccatggtgtt 120

Qy     164 gctgctgtcctccttctccgatcaatgcaatgtccc 199
           ||||||||| ||||||||||||||   ||  | |  ||
Db     121 gctgctgtactccttctccgatgcctgtgaggagcc 156


RESULT  15
US-08-435-149-17
; Sequence 17, Application US/08435149
; Patent No. 5866402
;  GENERAL INFORMATION:
;    APPLICANT:  INNIS, MICHAEL A.
;    APPLICANT:  ZAROR, ISABEL
;    APPLICANT:  CREASEY, ABLA A.
;    TITLE OF INVENTION:  CHIMERIC MCP AND DAF PROTEINS WITH CELL
;    TITLE OF INVENTION:  SURFACE LOCALIZING DOMAIN
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CHIRON CORPORATION
;      STREET:  INTELLECTUAL PROPERTY - R440, P.O. BOX 8097
;      CITY:  EMERYVILLE
;      STATE:  CALIFORNIA
;      COUNTRY:  U.S.A.
;      ZIP:  94662-8097
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/435,149
;      FILING DATE:  05-MAY-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SAVEREIDE, PAUL B.
;      REGISTRATION NUMBER:  36,914
;      REFERENCE/DOCKET NUMBER:  0989.001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (510) 601-2585
;      TELEFAX:  (510) 655-3542
;      TELEX:  N/A
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1878 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-435-149-17


  Query Match            18.6%;  Score 116.6;  DB 2;  Length 1878;
  Best Local Similarity   84.5%;  Pred. No. 1.1e-27;
  Matches  131;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;
```

```
Qy      45 ctgctcacctccggataaatcacggggtctcccgcgccgctcatggcgcctcccgtccgt 104
           |||||  |||||||| |||| || | |||| ||||||||| ||||| |||||||| |||
Db       1 CTGCTTTCCTCCGGAGAAATAACAGCGTCTTCCGCGCCGCGCATGGAGCCTCCCGGCCGC 60

Qy     105 ctcgagcgtccctttccttcccggcgctttcctgggttgcttctggcggccctggtgttg 164
           | |||| |||||||||||||| ||||||||||||||||||||||||||||| ||||||||
Db      61 CGCGAGTGTCCCTTTCCTTCCTGGCGCTTTCCTGGGTTGCTTCTGGCGGCCATGGTGTTG 120

Qy     165 ctgctgtcctccttctccgatcaatgcaatgtccc 199
           ||||||| ||||||||||||||   ||  | |  ||
Db     121 CTGCTGTACTCCTTCTCCGATGCCTGTGAGGAGCC 155


Search completed: October  9, 2002, 16:39:24
Job time: 4961 sec
```